CHARLES UNIVERSITY IN PRAGUE

Faculty of Science

Developmental and Cell Biology

The role of vent genes family in early development and brain development

Úloha genové rodiny vent v časném embryonálním vývoji a ve vývoji mozku

Ph.D. Thesis

Peter Fabian

Supervisor: Zbynek Kozmik, Ph.D.

Department of Transcriptional regulation

Institute of Molecular Genetics of the ASCR, v. v. i.

Prague, 2016

I DECLARATION

I declare that I wrote the thesis independently and that I cited all the information sources and literature. This work was not presented to obtain another academic degree or equivalent.

Prague 2016 Peter Fabian

ACKNOWLEDGMETNS

First and foremost, I would like to express my sincerest gratitude to Dr. Zbynek Kozmik for his supervision, guidance, and encouragement. This thesis would not have been possible without his support, great kindness, and patience.

I also wish to thank my colleagues, especially Chrysa Pantzartzi, Jan Masek, Iva Dobiasovska and Ira Kozmikova and all members from the Department of Transcriptional Regulation. It was really pleasure to work with all of them.

Finally, it is a pleasure to thank my family and especially my wife Zuzana.

CONTENTS

II ABBREVIATIONS

A/P	antero-posterior
ActR	activing receptor
bmp	bone morphogenetic protein
BMPR	bone morphogenetic protein receptor
boz	bozozok
CNS	central nervous system
chd	chordin
dkk	dickkopf
D/V	dorso-ventral
en	engrailed
fgf	fibroblast growth factor
gbx	gastrulation brain homeobox
GRN	gene regulatory network
HB	homeobox
MHB	midbrain-hindbrain boundary
MAPK	mitogen-activated protein kinase
otx	orthodenticle homeobox
pax	paired box
TF	transcription factor
TGF-β	transforming growth factor-β

III ABSTRACT (ENGLISH)

In chordates, the central nervous system (CNS) is derived from the dorsal part of gastrula. Induced dorsal part of the embryo - the neural plate - gives rise to the neural tube or primordial brain.

The developing dorsal part of the embryo is shaped by BMP/Smad signaling from the ventral part. Using the basal chordate amphioxus, we show here the conserved evolutionary role BMP/Smad signaling in axial cell fate determination. Pharmalogical inhibition of BMP/Smad signaling induces dorsalization of *Branchiostoma floridae* (amphioxus) and *Oryzias latipes* (medaka) embryos and expansion of neural plate markers. We provide evidence for the presence of the positive regulatory loop within the BMP/Smad signaling network of amphioxus. Thus, our data suggest that early emergence of a positive feedback loop within the BMP/Smad signaling network may represent a crucial molecular event in the evolutionary history of the chordate cell fate determination.

The dorso-ventral body axis formation is mediated by genes of the vent family, which are the direct targets of BMP/Smad signaling. The function of vent gene family in early development is relatively well known, however, its role in developing CNS is not yet clear. Therefore, we decided to manipulate *vox* transcription factor, a vent family member. Using medaka fish model, we show here the role of *vox*, in developing brain. The loss of function of *vox* does not lead to developmental defects, likely due to functional redundancy. Overexpression of *vox* shows that it interferes with the gene regulatory network during the maintenance stage of developing MHB. Collectively our data suggest that genes from the vent family, in addition to their crucial role in body axis formation, may play a role in regionalizing of vertebrate CNS.

IV ABSTRAKT (CZECH)

Centrální nervová soustava (CNS) je u všech strunatců odvozena z hřbetní strany vyvíjejícího se embrya. Postupným vlivem signalizace vzniká z hřbetní strany neurální ploténka, ze které se následně vyvine budoucí mozek spolu s neurální trubicí.

Budoucí CNS je formována účinkem BMP/Smad signalizační dráhy z břišní části embrya směrem k hřbetní. Konzervovanou roli této BMP/Smad signalizační dráhy v průběhu evoluce zde ukazujeme potvrzením jejího vlivu při stanovení buněčného osudu v rámci tělní osy i v případě bazálního strunatce *Branchiostoma floridae* (kopinatce). U kopinatce vede farmakologická inhibice BMP/Smad signalizační dráhy k rozšíření exprese genů typických pro neurální ploténku a k dorzalizaci embrya, stejně tak jako u ryby *Oryzias latipes* (medaka). Poskytujeme důkaz přítomnosti pozitivní autoregulace BMP/Smad signalizační dráhy u kopinatce. Zároveň naše data naznačují možnost časného vzniku pozitivní autoregulace BMP/Smad signalizační dráhy jako klíčové události v evoluční historii stanovení buněčného osudu u strunatců.

Tvorba dorso-ventrální osy těla je zprostředkovány geny genové rodiny vent, které jsou přímým cílem BMP/Smad signalizace. Ačkoliv je funkce genů genové rodiny vent jako transpčních faktorů v rámci časného embryonálního vývoje poměrně dobře popsána, jejich role v rámci tvorby CNS doposud není jasná. Pro osvětlení této role jsme se rozhodli pro manipulaci člena vent genové rodiny – genu *vox*. Prostřednictvím rybího modelu zde ukazujeme roli *vox* ve vývoji mozku.

Ztráta funkce genu *vox* nevede k vývojovým defektům, pravděpodobně díky redundanci genetické informace. Naopak zvýšená exprese genu *vox* v tzv.isthmu mezi středním a zadním mozkem v průběhu raného vývoje, ukazuje propojení s genovou regulační sítí. Společně naše data naznačují roli genů z genové rodiny vent při regionalizaci CNS jako další roli navíc k jejich zásadní úloze v rámci tvorby tělní osy.

V AIMS OF THE STUDY

During early embryonic development, the embryo is becoming asymmetrical and needs to distinguish position of body axis. In the gastrula-stage, the match between the dorsal and ventral sides is essential for patterning of tissues. To distinguish ventral and dorsal part of the body, the embryo requires BMP-like morphogens on ventral side, which antagonized the dorsal side of the embryo. This process is also mediated, by vent gene family. Although the role of BMP/Smad signalling and vent genes family in early development has been studied intensively in many model organisms, there is lack of data about the situation in basal chordates (e.g. *Branchiostoma floridae*).

Vent family genes are transcription factors that bind DNA via homeobox domain and regulate expression of other genes. Their unique role during early development is relatively well known. Moreover, the new roles (e.g. in hematopoiesis or cancer) of vent family genes are emerging. *vox*, a member of vent family, helps specify non-neural ectoderm in teleost fish, however, its role during brain segmentation has never been described. Therefore, we manipulated *vox* gene in the teleost fish *Oryzias latipes* - medaka - to elucidate its function in one of the crucial signaling centres of developing brain, the isthmic organizer or midbrain-hindbrain boundary (MHB). MHB is a highly conserved vertebrate signaling center, acting to pattern and establish neural identities within the brain.

Specific aims of my PhD thesis were the following:

1. to investigate the evolution of BMP/Smad signaling pathway function during embryogenesis in *Branchiostoma floridae* (amphioxus).

2. to characterize the role of *vox*, a vent gene family member, in the developing brain;

VI INTRODUCTION

VI.1. BMP/Smad signalling

Bone morphogenetic proteins (BMPs) belong to transforming growth factor-β (TGF-β) of superfamily ligands, which play important role in a myriad of biological activities (Bier and De Robertis 2015) such as pattern formation, morphogenesis, organogenesis, cell differentiation, homeostasis, regulation of iron and energy metabolism (Bier and De Robertis 2015; Reddi and Reddi 2009). The signaling is initiated by the interaction of BMP ligands with their heterodimeric receptors and elicits a variety of intracellular responses. BMPs signal via both the canonical, Smad-dependent pathway and various non-canonical pathways (Wang et al. 2014). In the canonical BMP Smad-dependent pathway, the BMP-receptor (BMPR) interaction results in the activation of Smad-mediated gene expression. In addition to Smads, BMPs also activate non-Smad (non-canonical) signalings, namely the mitogen-activated protein kinase (MAPK), influencing a family of molecules including p38 and ERK1/2 (Cao and Chen 2005; Reilly et al. 2005). Nevertheless, due to the complexity of BMP/Smad signaling and its broad engagement in various biological events, there are still many things that are not clarified.

Canonical BMP/Smad signaling

BMPs were originally discovered as extremely important morphogens due to their ability to activate the formation of cartilage and bone. BMPs have been extensively studied from 1960's (Luyten et al. 1989; Urist 1965; Wozney 1992; Wozney et al. 1988) and to date, over 20 BMP family members have been identified and characterized (reviewed Bier and De Robertis (2015)). BMPs are now regarded as a group of pivotal morphogenetic signals that orchestrate tissue architecture all over the body.

In the cell, the BMP molecules are first expressed as large precursors. These large molecules are then dissected at a dibasic site, and the C-terminal part with active domain is released. Before secretion, BMPs contain a signal peptide, pro-domain and mature peptide. Subsequent cleavage of the signal peptide, the precursor

protein undergoes glycosylation and dimerization. The mature bioactive dimeric BMP is secreted out of the cell and the pro-domain is cleaved by proteolytical cleavage. BMPs are secreted as either heterodimers or homodimers (Luyten et al. 1989).

BMP signal is transduced into the cell by a heterotetrameric receptor complex (Fig.1), which consists of the serine/threonine kinase receptors type II and type I. BMPs can bind to each of the three subtypes of receptor type II, as well as the three subtypes of receptor type I. The cross-talk between the both BMP receptors (type II and type I) is indispensable for signal transduction. Three type II receptors for BMPs have been identified: BMPR-II and the activing receptors ActR-II and ActR-IIB (Bier and De Robertis 2015; Kawabata et al. 1995; Nohno et al. 1995; Yamashita et al. 1995). Also, three type I receptors have been identified for BMP binding, namely BMPR-IA or ALK-3, BMPR-IB or ALK-6 and the activin receptor ActR-IA or ALK-2 (Koenig et al. 1994; Macias-Silva et al. 1998; ten Dijke et al. 1994). While BMPR-IA, IB, II are BMP-specific, ActR-IA, II, and IIB receptors are also signaling for activin (another TGF-β ligand). The expression of these receptors varies in different tissues. When BMP dimers bind the extracellular domain of the type II receptors they induce a conformational change. The serine/threonine kinase domains of type II receptors, which are constitutively active, phosphorylate type I receptors at the glycine-serine domains that leads to the activation of type I receptor kinases.

The heterotetrameric-activated receptor complex consists of two pairs of type II and type I receptors (Bier and De Robertis 2015; Moustakas and Heldin 2002). This leads to the involvement of the pathway restricted intracellular proteins - Smads. Smad1/5/8 directly and transiently interacts with activated BMP receptors type I, which phosphorylate Smads at the C-terminal SSXS motif (Chen et al. 1997; Hoodless et al. 1996; Nishimura et al. 1998). Phosphorylated Smads are subsequently released from the receptor and physically interact with Smad4 to form a heteromeric complex. Smad4 is a common Smad mediator and acts as a shared partner. This heteromeric Smad complex is translocated into the nucleus where triggers the transcription of specific target genes (Heldin et al. 1997). Smad1/5 can also bind to DNA directly – without Smad4, however, the affinity is relatively low. For direct DNA binding, the interaction of Smad1/5 with sequence-specific DNA binding proteins is critical (Derynck et al. 1998). It has been reported that Smad1/5 interacts with the transcription factor (TF) Runx2 (Hanai et al. 1999; Lee et al. 2000; Moon et al. 2014; Zhao et al.

2003) and activates the transcription of target genes (Chikazu et al. 2002; Leboy et al. 2001).

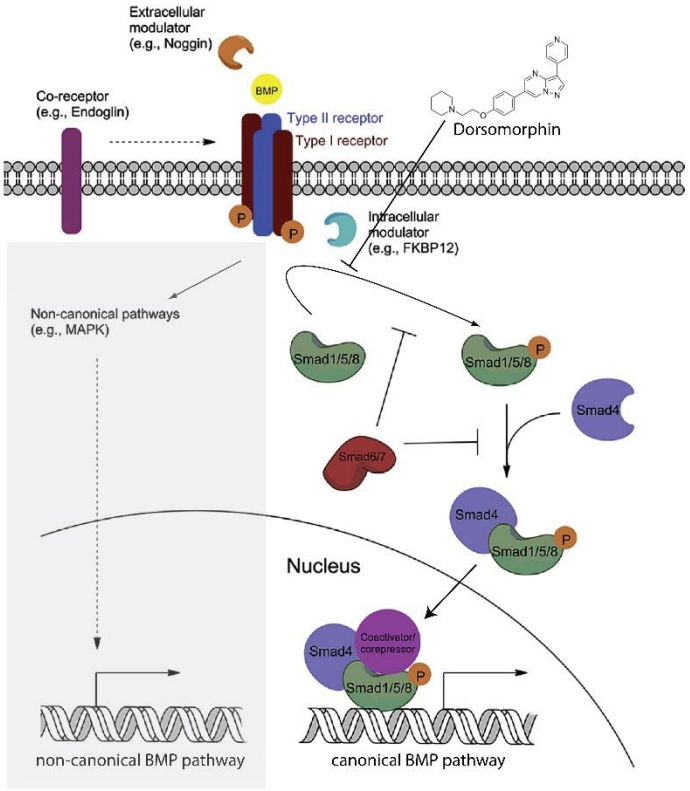


Fig.1. Model of BMPs signaling. BMPs initiate the signal transduction cascade by binding to type II and I receptors and forming activated heterotetrameric complex. Left: In non-canonical BMP signaling: various pathways, including the MAPK cascade, can lead to regulation of gene expression. Right: In the canonical pathway, and the type I receptor phosphorylates the Smad1/5/8 which then associates with the Smad4. The heterodimeric Smad complex translocates to the nucleus, to regulate gene expression. BMP signaling is modulated in the nucleus (with co-activators or co-repressors) extracellularly (e.g., Noggin), intracellularly (e.g., FKBP12, microRNAs, phosphatases, and I-Smads), and by co-receptors in the plasma membrane (e.g., Endoglin). Dorsomorphin inhibits BMP signaling through the Smad pathway, likely by affecting BMP receptor type I kinase activity (adopted from Wang et al. (2014)).

BMP/Smad signaling is modulated by diverse proteins at distinct points and contains positive/negative feedback loops. At the level of initiation of the signaling cascade, secreted antagonists, such as Cerberus, Dan, Gremlin and Noggin, specifically bind BMPs and form extracellular complexes. These protein inhibitors act

upstream from receptors and block both Smad-dependent and Smad-independent signaling.

In the intracellular compartment, the signal can be modulated by the activation of inhibitory proteins such as Smurf1 or Smad6/7. In the nucleus, there are many co-activators needed for the activation of specific target genes, whose activity can be inhibited by co-repressors (Kawabata et al. 1998).

Overall BMP/Smad signaling is involved in many biological processes, one of which is its fundamental role in early body axis formation.

VI.2. Dorsoventral body axis formation

During early embryonic development, embryo is becoming asymmetrical and needs to distinguish position of body axis. To distinguish ventral and dorsal part of the body, the embryo requires on one side BMP-like morphogens, and on the other side their inhibitors. Dorsal fate is supported via BMP inhibitors that act in an overlapping manner (Khokha et al. 2005). On the embryonic ventral side, the BMP molecules are expressed and work together to promote ventral fates (Reversade et al. 2005). The match between the dorsal and ventral sides, where each is antagonizing the other, is essential for patterning of tissues in the gastrula-stage. BMP signaling triggers transcription of downstream genes via activation of Smad proteins. But how can transcriptional activators such as BMPs downregulate the expression of genes at dorsal side? BMP signaling activate a set of transcriptional repressors, members of the vent gene family. Vent family genes inhibit the expression of the organizer genes encoding the BMP inhibitors. On the contrary, BMP inhibitors, secreted dorsally, prohibit the BMPs from inducing ventral fate by blocking the spherical interaction of BMPs with their receptors. This cross-talk creates gradient of BMP signaling from ventral to the dorsal sides of the embryo. Moreover this distinction is crucial for subdividing different fates of the germ layers.

The upper part of embryonic body is subdivided into neuroectoderm on the dorsal side and epidermis on the ventral side. From the equatorial part of the embryo, the dorsal side is predetermined to differentiate into head mesoderm and notochord, whereas, the ventral side gives rise to somites (Kimelman and Pyati 2005).

Dorsoventral patterning

Establishment of dorso-ventral (D/V) axis in vertebrates involves a number of factors: signaling molecules (i.e., TGF-β and Wnt factors), their protein antogonists e.g chordin (chd), dickkopf (dkk) and TFs (e.g. vent family genes and β–catenin) (Glinka et al. 1998; Glinka et al. 1997; Hashimoto et al. 1999; Larabell et al. 1997; Van Hul et al. 2002). In *Xenopus* and *Danio rerio* (zebrafish), the very first asymmetry is organised by the unequal distribution of β–catenin molecules. Increased concentration of β–catenin at the dorsal side is caused by its translocation and partially unknown determinants from the vegetal pole (Jesuthasan and Stahle 1997; Larabell et

al. 1997; Mizuno et al. 1999; Sokol 2015). The action of β–catenin as an important transcription cofactor introduces the (D/V) body axis specification. The initial expression of dorsal-specific genes, such as *bozozok/dharma* (*boz*) in zebrafish, is triggered by the cooperation of HMG-box factor Tcf/Lef with β–catenin. *boz*, directly co-regulated by β–catenin, acts as transcriptional repressor and represses several genes that help to specify the ventral part of the body (Bier and De Robertis 2015; Fekany et al. 1999; Koos and Ho 1999; Yamanaka et al. 1998). This fact has been supported by the analysis of a zebrafish *boz* mutant that possesses defected dorsoanterior parts of the body. It has been reported that this factor represses BMP and Wnt signaling in the anterior neuroectoderm (Solnica-Krezel and Driever 2001), whereas the ventralizing activity of the body is triggered by TGF-β and especially, BMPs family (Balemans and Van Hul 2002). The ventralizing potential of *bmp2*, *bmp4*, or *bmp7 and* of their transducers, *smad1* and *smad5*, was confirmed by the gain of function approach in *Xenopus laevis*, (Piccolo et al. 1996; Reversade et al. 2005; Wessely et al. 2004); on the other hand, depletion of BMP signaling determines a dorsalizing effect (Fig.2). Similar studies confirmed the role of BMP/Smad signaling using zebrafish various mutants of the (*bmp2, bmp7, smad5, alk8*) (Bauer et al. 2001; Dick et al. 2000; Hild et al. 1999; Kishimoto et al. 1997; Kodjabachian et al. 1999; Martinez-Barbera et al. 1997; Mullins et al. 1996; Nguyen et al. 1998; Schmid et al. 2000; Tuazon and Mullins 2015).

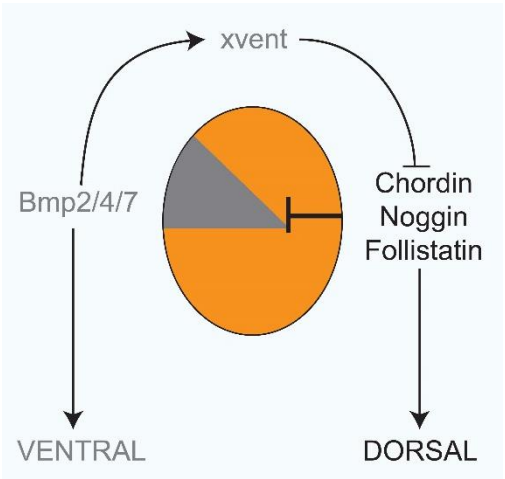


Fig.2. BMP signaling establishes a self-regulating morphogenetic field. Simplified model for the creation of a ventral-dorsal BMP gradient (indicated by the grey triangle). Ventral BMPs and dorsal BMP inhibitors mutually inhibit each other (adopted from Kimelman and Pyati (2005)).

Spemann organizer is a secreting centre of dorsalizing factors, such as Noggin, Follistatin, and Chd. These factors bind and block the ventralizing activity of BMP proteins (Piccolo et al. 1996; Sasai et al. 1994; Zimmerman et al. 1996). The activity of chd is tightly regulated by the metalloprotease Tolloid that cleaves Chd protein, thereby indirectly promoting BMP signaling (Berry et al. 2010). The role of *chd* has been studied using zebrafish *chordino* mutant, in which the ventralizing effect is demonstrated as an enlarged ventral region at the expense of the head and the anterior part of the trunk (Gonzalez et al. 2000; Hammerschmidt et al. 1996; Mullins 1999; Schier 2001; Schulte-Merker et al. 1997). At the end of gastrulation, the embryo is subdivided to the ventral and dorsal side. The embryonic neural tube and later the central nervous system will be formed from the dorsal side.

VI.3. Early brain development and midbrain-hindbrain boundary

The central nervous system (CNS) arises from the dorsal side of the vertebrate gastrula. It is induced by a combination of signals that arise from the posterior margin of the embryo and later from the gastrula organizer. Subsequently, neural tube is formed from the induced neural plate (reviewed in Wurst and Bally-Cuif (2001)). This tube is subdivided along its antero-posterior (A/P) axis into four major divisions: the forebrain, the midbrain, the hindbrain and the spinal cord. “Classical” experiments, which consisted of cutting and pasting different pieces of the neural tube in different orientations and locations, had identified two such signaling centres. One of them is the anterior neural ridge, which lies at the junction between the prospective forebrain and the anterior ectoderm, and is necessary for the preservation of forebrain. The second signaling centre is midbrain-hindbrain boundary (MHB). The spatio-temporal expression patterns of MHB genes seem to be broadly conserved throughout evolution. Slight differences exist among species, notably when comparing the onset of expression of the different genes (e.g. *pax2* or *fgf8*). MHB lies at the junction between the midbrain and hindbrain, and is necessary and sufficient for the development of mesencephalic and metencephalic structures. Formation of the MHB organizer is traditionally thought of as comprising three stages, namely positioning, induction and maintenance (reviewed in Wurst and Bally-Cuif (2001) and Dworkin and Jane (2013)).

Midbrain-hindbrain boundary

Patterning of the primordial brain begins shortly after gastrulation, when the neural plate begins to form. During the neurulation process of teleost fish, the neural plate firstly transforms into a transient structure called neural keel, which later gives rise to the neural rod (approximately at 5-somite stage). Finally, the neural tube forms at 10-somite stage. The neural keel is regionalized both across the D/V and the (A/P) axis. Later on, the neural rod is divided into three primary brain regions: the forebrain, the midbrain and the hindbrain, as a result of intensive signaling cross-talk. D/V patterning of the neural rod (neural tube) is controlled via ventralizing signals – SHH from notochord cells - and via dorsalizing signals – mainly BMP4 from epidermal ectoderm (Echelard et al. 1993; Lee and Jessell 1999). A/P patterning, on the other hand, is organized by local signaling centers (Echevarria et al. 2003; Jessell and Sanes 2000). One of them is a group of cells at the MHB, also called the isthmic organizer (Houart et al. 1998; Wilson and Houart 2004; Wurst and Bally-Cuif 2001). MHB is both required and sufficient for the cell fate induction of the surrounding tissues, which are later developed into mesencephalic and metencephalic regions (Liu and Joyner 2001; Raible and Brand 2004; Wurst and Bally-Cuif 2001). This junction is established, induced and maintained by the expression of several TFs, most notably *otx*, *gbx*, *pax* and *en*, and signaling molecules from Wnt and Fgf families. These factors have been found in all studied vertebrate species and are often regarded as the core MHB cascade (reviewed in Dworkin and Jane (2013) and Wurst and Bally-Cuif (2001)).

In vertebrates, the establishment of MHB is determined at the borders of the expression domains of homeobox (HB) transcription repressors from the otx (orthodenticle homologue) and gbx (gastrulation brain HB) families. In zebrafish, MHB positioning is specified at the interface of the expression domains of HB TFs *otx2* (presumptive midbrain) and *gbx1* (presumptive hindbrain) (Rhinn et al. 2009; Su et al. 2014). Positioning of the MHB primordium (*otx2*/*gbx1* borders) in the zebrafish neural plate is mediated by Wnt signaling from the precursors of lateral mesendoderm (Rhinn et al. 2009; Rhinn et al. 2005), and not the primordial notochord - axial mesendoderm - as it was thought before. One of the critical molecules of the developing MHB is *fgf8*. The onset of *fgf8* expression is triggered during early neural plate at the presumptive MHB region. *fgf8* enhances the expression of *gbx1*/2, which

then repress the activity of *otx2* (Liu and Joyner 2001; Sato and Joyner 2009). Most interestingly, Fgf8 can be still detected within the neuroectoderm in the absence of all these factors, indicating that *otx/gbx1/2* border rather than inductive region, may serve as re-shaper of the *fgf8* expression (Su et al. 2014). Once the expression of *otx2/gbx1* is established, the induction step of MHB with three parallel signaling pathways involving *pax*, *wnt* and *fgf* is triggered (Fig.3) (Canning et al. 2007; Lun and Brand 1998; Rhinn and Brand 2001).

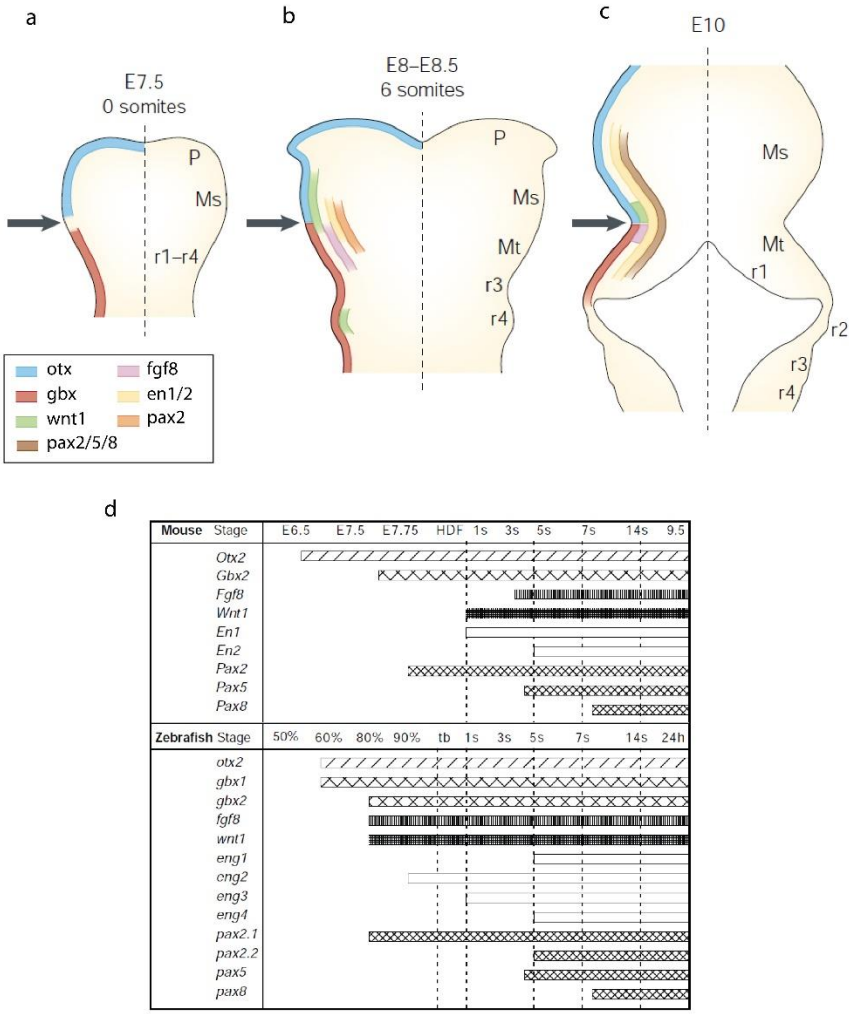


Fig.3. Gene expression patterns at the MHB. (**a-c**) Dorsal view of developing mouse brain, anterior to the top. (**d**) The onset of expression MHB genes in mouse and zebrafish. Ms, mesencephalon; Mt, metencephalon; P, prosencephalon; r, rhombomeres; symmetry axis - dashed line. MHB marked with an arrow. (adopted fromWurst and Bally-Cuif (2001) and Rhinn and Brand (2001)).

Expression of these genes is also called core-MHB cascade.The secreted factor Wnt1 is first expressed largely across the entire midbrain before it becomes restricted to the *otx2*-positive territory (Bally-Cuif et al. 1995), whereas *Pax2* expression crosses the *otx2*/*gbx2* border (Rowitch and McMahon 1995).

Experiments in the chicken, using beat or graft transplantation approach, indicated that Fgf8 possesses inductive potential. Ectopic presence of Fgf8 induces the MHB cascade and organizes surrounding tissues, leading to the formation of both the tectum and the cerebellum (Crossley et al. 1996). However, zebrafish lacking *fgf8* do not form cerebellum (Reifers et al. 1998). These data show that the gradients of Fgf8 expression must be critically regulated. Nevertheless, the core-MHB cascade is still initiated in animals lacking Fgf8 (Chi et al. 2003; Reifers et al. 1998), although the expression of core-MHB genes is not subsequently maintained. Slightly later, *en2, en3 and en1*, as well as *pax2.2, pax5* and *pax8* are expressed across MHB (Lun and Brand 1998). After the initiation phase, *fgf8* expression is positively regulated by *gbx2*, while *otx2* and *gbx2/fgf8* regulate each other negatively, and these interactions lead to the establishment and maintenance of sharp expression borders (Rhinn et al. 2003). Concomitantly, the expression regions of *fgf8*, *wnt1*, *en1/2/3*, *pax2.1/2.2/5/8* become interdependent and establish a positive regulatory loop that is required for maintaining the midbrain-hindbrain identity (Fig.4) (Dworkin and Jane 2013; Wittmann et al. 2009; Wurst and Bally-Cuif 2001).

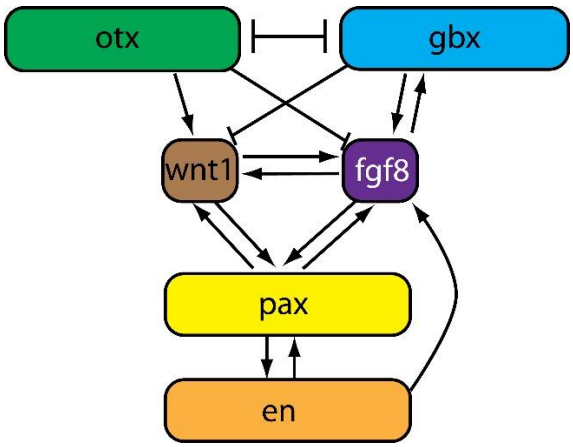


Fig.4. Scheme of MHB genes interactions. Arrowheads represent positive regulation. Bars represent negative regulation (adopted from Wittmann et al. (2009)).

The maintenance of MHB relies on continuous interactions of the core MHB factors, namely *fgf8*, *wnt1*, *en1/2*, *pax2.1* (Dworkin and Jane 2013; Joyner et al. 2000; Rhinn and Brand 2001; Wurst and Bally-Cuif 2001). Perturbation of any gene from this GRN has no crucial effect on the induction, however, it will lead to severe defects in the maintenance of the isthmic organizer (Dworkin and Jane 2013; Joyner et al. 2000; Rhinn and Brand 2001; Wurst and Bally-Cuif 2001). Other TFs (pou2, Brn1, Sef, Tapp1, etc.), various signaling pathways and non-coding RNAs have also been discovered to operate upstream of the MHB cascade. The regulation of MHB cascade seems to be more complex and also the function of large number of genes that have been detected at the MHB (Thisse et al. 2004) remains unknown.

VI.4. Vent family genes

VI.4.1. Overview

Genes belonging to the Vent family encode TFs that belong to NK-like class, from the antennapedia superclass, which is a subgroup of Homeodomain superfamily. A primary function of HB proteins is transcriptional regulation of other genes during development and differentiation. HB TFs from the vent gene family are known as important players in the formation of the ventral mesoderm during early development, whereas present work on vent family genes is suggesting new roles in endoderm specification, hematopoiesis even in maturation of dendritic cells and malignant myelopoiesis.

VI.4.2. Vent family genes as members of homeodomain superfamily

The homeodomain or HB is a DNA binding domain, hence HB proteins are essentially TFs. They have been shown to play major role in many developmental processes of animals, as well as of fungi and plants. A primary function of HB proteins is to regulate the expression of other genes in development and differentiation. The HB is described as a conserved DNA motif of approximately 180 base pairs and encodes the about 60 amino acid long (Larroux et al. 2007). The first genes found to encode HB proteins were homeotic genes from *Drosophila*, from which the name "homeo" box was derived (Burglin 2011; Laughon and Scott 1984). Nevertheless, not all HB genes are homeotic genes, and not all homeotic genes are HB genes (Burglin 2011). Vent family genes belong into antennapedia HB superclass. In antennapedia superclass, a number of genes are organized into gene clusters, i.e. the HOX cluster, and the NK cluster, however, many of genes are referred to as "dispersed". These genes do not encode large conserved domains outside of the HB, but only small motifs. Vent family genes belong intodispersed NK-like class with families like BarHl, Barx, Bsx, Dbx, Dlx, Emx, possibly En, Hhex, Hlx, Nanog, Nkx2.1, Nkx2.2, Nk6, and Vax (Burglin 2011).

VI.4.3. Vent genes family in dorsoventral patterning

Vent genes family play role in early development and differentiation of embryonic body. Body axis formation is regulated by the orchestrated action of different TFs and signaling molecules. BMP signaling pathway plays crucial role in the control of D/V patterning of every bilateral embryo. This process is mediated, by the vent-like genes (Imai et al. 2001; Kawahara et al. 2000a; Kawahara et al. 2000b; Shimizu et al. 2002).

The pioneer experiments were done on *Xenopus*, however, the latest studies have been performed on zebrafish. Zebrafish possesses three vent family genes, *vox* (*vega1*), *vent* (*vega2*) and *ved*, which encode transcriptional repressors implicated in patterning the D/V axis as well (Gilardelli et al. 2004; Imai et al. 2001; Kawahara et al. 2000a; Kawahara et al. 2000b; Melby et al. 2000; Shimizu et al. 2002). *ved* has a maternal expression and the onset of embryonic expression of *vox, vent and ved* is triggered by the maternally expressed TF Runx2b (Flores et al. 2008). Another gene that is involved in the direct activation of *vox at* late blastula and gastrula stages is the TF *Pou5f1* (Belting et al. 2011). After zygotic genome activation, *vox* is expressed in all blastomeres and, at dome stage, is down-regulated in the dorsal cells and in the yolk syncytial layer cells. *vent* is activated later, at sphere stage (late blastula), in the ventrolateral region (Kawahara et al. 2000b; Melby et al. 2000). The expression is later maintained by Wnts until the mid-gastrula stage (Ramel and Lekven 2004; Varga et al. 2007). The positive autoregulatory loop between BMP pathway and *vox*/*vent* genes is established at mid-gastrula stage (Imai et al. 2001; Kawahara et al. 2000a; Kawahara et al. 2000b; Melby et al. 2000; Ramel and Lekven 2004), whereas at earlier stages *vox*, *vent* and *ved* are relatively independent of BMP signaling. *ved, vox and vent* are negatively regulated by *bozozok*/*dharma and* inhibit expression of the organizer genes. In addition, loss of function experiments have shown that these vent family genes have redundant functions in the organization of the dorsal region of the embryo (Shimizu et al. 2002). Possible downstream genes are dorsal genes, such as *dkk1*, *squint*, *cyclops*, *floating head*, *chd* and *goosecoid* (Kawahara et al. 2000a; Kawahara et al. 2000b; Melby et al. 2000). Although *vent* and *vox* can control analogous or same target genes, differences in the activities of these vent family genes in overexpression assays indicate that these TFs may have different regulatory functions (Kawahara et al. 2000a; Kawahara et al. 2000b; Melby et al. 2000).

Vent family genes, among which *vox*, belong to a group of HB TFs that play a significant role during early development. Overexpression of these genes leads to strong ventralization of zebrafish embryos (Gilardelli et al. 2004). *vox and* other members of the vent family act as transcriptional repressors (Gilardelli et al. 2004; Kozmikova et al. 2013; Zhao et al. 2013). The role of the *vox* gene (or vent family genes) has been relatively well studied during early development, nevertheless, it is largely unknown during later developmental stages.

VI.4.4. New roles of vent family genes

Recently, VentX, the human homolog of vent family genes, has been located in the syntenic regions of dog and chicken genomes, but not of mouse and rat, probably due to loss of HB genes in the ancestry of rodents, rather than in the ancestry of human (Zhong and Holland 2011). VentX has been defined as a novel hematopoietic TF controlling differentiation and proliferation of hematopoietic and immune cells (Gao et al. 2007; Gao et al. 2012; Wu et al. 2011a). Initially identified as a novel Tcf/Lef associated antagonist of the canonical Wnt signaling in *Xenopus* development, VentX was found to be an activator of the p21/p53 expression and p16ink4a/Rb tumor suppression pathways (Gao et al. 2010; Wu et al. 2011b). The critical role of VentX was indicated also in the development of hematopoietic cells where plays role in controlling the proliferation and differentiation of CD34 cells as well as monocytes to macrophage terminal differentiation (Gao et al. 2012; Wu et al. 2011a). Latest data also suggest that VentX might be a key regulator of dendritic cells (antigen presenting cells) differentiation and maturation (Wu et al. 2014).

VII RESULTS AND DISCUSSIONS

VII.1. Comments on presented publications

Kozmikova I, Candiani S, **Fabian P**, Gurska D, Kozmik Z.
Essential role of Bmp signaling and its positive feedback loop in the early cell fate evolution of chordates.
Dev Biol. 2013 Oct 15;382(2):538-54.

During development, early separation of cell fate domains in chordates occurs prior to the final specification of ectoderm to neural and non-neural, as well as of mesoderm to dorsal and ventral part. Maintenance of such division through the establishment of an exact border between the domains is required for the formation of highly differentiated structures such as notochord or neural tube.

We hypothesized that the key condition for efficient cell fate separation in the chordate embryo is the presence of a positive feedback loop for BMP/Smad signaling within the GRN that underlies early axial patterning. We wanted to investigate the evolution of BMP/Smad signaling pathway function during embryogenesis at the invertebrate-to-vertebrate transition, therefore we performed a study of the role of this pathway at embryonic stages in *Branchiostoma floridae* (amphioxus). Amphioxus is a suitable model organism to address evolutionary questions of early embryological development because of its unique position in the chordate tree. Moreover, amphioxus genomic, morphological, and developmental attributes are probably deeply analogous to those of the chordate ancestor. Here, we investigated the role of BMP/Smad signaling in axial cell fate determination in amphioxus. Pharmacological inhibition of BMP/Smad signaling induces dorsalization of both *Oryzias latipes* (medaka) and amphioxus embryos and expansion of neural plate markers. This is consistent with an ancestral role of BMP/Smad signaling in chordate body axis patterning and neural plate formation. Furthermore, we provide confirmation for the presence of the positive regulatory loop within the BMP/Smad signaling network of amphioxus. Using mRNA microinjections, we found in contrast with vertebrate Vent genes, which trigger the expression of Bmp4, amphioxus Vent1 is probably not responsible for activation of the cephalochordate ortholog Bmp2/4. Cis-regulatory analysis

of amphioxus Bmp2/4, Admp and Chd promoters in medaka fish embryos revealed notable conservation of the gene regulatory information between basal chordates and vertebrates.

Our data suggest that appearance of a positive regulatory loop within the BMP/Smad signaling network may represent a crucial molecular event in the evolution of the chordate cell fate determination.

Author's contribution: I did the cloning of *admp* gene, RNA *in situ* hybridization on medaka fish embryos, as well as inhibition of BMP pathway using dorsomorphin on medaka.

Fabian P, Pantzartzi CN, Kozmikova I, Kozmik Z.
vox homeobox gene: a novel regulator of midbrain-hindbrain boundary development in medaka fish.
Dev Genes Evol. 2016

Induction of the neural plate is considered the initial step in the development of the CNS. Patterning of the primordial brain begins shortly after gastrulation, when the neural plate begins to form. The MHB is one of the key organizing centers of the vertebrate CNS. Its patterning is governed by a well-described gene regulatory network GRN, involving several TFs, namely *pax*, *gbx*, *en* and *otx*, together with signaling molecules of the Wnt and Fgf families.

Here we describe the onset of these markers in medaka early brain development in comparison to previously known *Danio rerio* (zebrafish) expression patterns. Moreover, we show for the first time that *vox*, a member of the vent gene family, is expressed in the developing neural tube similarly to CNS markers. The role of the *vox* gene (or vent family genes) has been relatively well studied during early development, nevertheless, it is largely unknown during later developmental stages. Our aim was to investigate the role of the *vox* HB gene in the early brain development. We specifically address the relationship between *vox* and the MHB program. We performed synteny and sequence conservation analysis to validate the identity of the medaka *vox* ortholog. We described its expression pattern during embryogenesis

and placed it into the framework of known brain markers. The loss of function of *vox* did not lead to developmental defects, likely due to functional redundancy. Using a heat-shock inducible *vox* line of medaka, we show that *vox* interferes with the GRN during the maintenance stage of developing MHB. Overexpression of *vox* leads to profound changes, most notably of *fgf8, a* crucial organizer molecule of MHB.

Our data suggest that genes from the vent family, in addition to their crucial role in body axis formation, may play a role in regionalizing of vertebrate CNS.

Author's contribution: I performed all the experiments, except the generation of the medaka heat-shock inducible *vox* line.

Fabian P, Kozmikova I, Kozmik Z, Pantzartzi CN.
Pax2/5/8 and *Pax6* alternative splicing events in basal chordates and vertebrates: a focus on paired box domain.
Front Genet. 2015 Jul 2;6:228.

Pax genes belong to a family of TFs, which contribute to embryonic development, and illustrate phylogenetically conserved functions, as evidenced by various animal models. The number of *Pax* homologs varies among species studied so far, due to gene and genome duplications which have affected Pax genes family. Pax genes have been identified in chordates, based on sequence similarity and functional domains, into four classes - namely Pax1/9, Pax2/5/8, Pax3/7, and Pax4/6. Many splicing events have been reported mainly for *Pax6* and *Pax2/5/8* genes. Of significant interest are those events that lead to Pax proteins with assumed novel properties, e.g. modified DNA-binding or transcriptional activity.

In the current study, a thorough analysis of *Pax2/5/8* splicing events from cephalochordates and vertebrates was performed. We focused more on *Pax2/5/8* splicing events in which the paired domain is involved. Three new splicing events were identified in medaka *Pax2.1 and Pax2.2* genes, one of which seems to be conserved in Acanthomorphata. Using representatives from protostome and deuterostome phyla, a comparative approach of the *Pax6* exon-intron structure of the paired domain was performed, during an attempt to calculate the time of appearance of the *Pax6(5a)*

isoform. As shown in our analysis, this splicing event is unique for Gnathostomata and is absent in the other chordate subphyla. Moreover, expressions of alternative spliced variants were compared between cephalochordates and fish species.

In summary, we identified three new splicing variants in medaka *Pax2* genes, one of which seems to be highly conserved in Acanthomorphata. We identified a re-occurring splicing event in medaka and zebrafish *Pax6* genes, generating the exon (5a) insertion. Our data indicate expansion of alternative splice variants in the paired box region of *Pax6*and *Pax2/5/8* genes during the course of vertebrate evolution

Author's contribution: I prepared RNA samples from defined stages of medaka and zebrafish development and performed the RT-PCR experiment.

Liegertová M, Pergner J, Kozmiková I, **Fabian P**, Pombinho AR, Strnad H, Pačes J, Vlček Č, Bartůněk P, Kozmik Z.
Cubozoan genome illuminates functional diversification of opsins and photoreceptor evolution.
Sci Rep. 2015 Jul 8;5:11885

Numerous animals sense light for vision and nonvisual photoreception. Light is captured by an eye which consists of opsin-based photopigment current in a photoreceptor cells. *Cnidaria* are probably the most basal phylum containing a camera type visual system. The evolutionary history of opsin genes in the animal kingdom has not yet been answered.

First, we have investigated the evolution of animal opsins by genome-wide approach of the box jellyfish *Tripedalia cystophora*, a cnidarian having complex lens-containing eyes and minor photoreceptors. A large number of opsin genes with distinct tissue- and stage- specific expression were identified. Our phylogenetic analysis classifies cubozoan opsins as a sister group to c-opsins and presents lineage-specific expansion of the opsin gene collection in the cubozoan genome. Functional analyses show evidence for the use of the Gs-cAMP signaling pathway in a small set of *T. cystophora* opsins, indicating the probability that the majority of other cubozoan opsins signal via distinct but not yet identified signaling pathway. Moreover, these

tests uncovered fine differences among individual opsins, suggesting possible fine-tuning for specific photoreceptor tasks. Moreover, the loss of phototactic behavior of living medusa, using pharmacological inhibition of the Gs-cAMP signaling pathway, unveiled its physiological role in vision.

Overall, our phylogenetic, expression, biochemical and behavioral analysis suggests that cnidarian intron-less opsins might have been derived from an ancient eumetazoan ciliary-like opsin possessing introns, by retro-transposition. Once docked in the genome, the ancient cnidopsin gene underwent multiplication, diversification and sensitivity modification. Individual opsins were thus accommodated for distinct functions in varied tissue photoreceptors - ocular, extraocular and larval. These opsins differ in stage- or tissue-specific expression, primary assembly and also in following cellular signaling – either via Gs-cAMP pathway or other G-protein pathways.

Author's contribution: I prepared the *Tripedalia cystophora* culture, I collected animal samples that were used for qRT-PCR experiments and performed the behavioral phototaxis test.

VII.2. Presented publications

VII.2.1. **Kozmikova I, Candiani S, Fabian P, Gurska D, Kozmik Z.** Essential role of Bmp signaling and its positive feedback loop in the early cell fate evolution of chordates. **Dev Biol. 2013 Oct 15;382(2):538-54.**

Developmental Biology 382 (2013) 538–554

Contents lists available at ScienceDirect

Developmental Biology

journal homepage: www.elsevier.com/locate/developmentalbiology

Evolution of Developmental Control Mechanisms

Essential role of Bmp signaling and its positive feedback loop in the early cell fate evolution of chordates

Iryna Kozmikova [a,1,*], Simona Candiani [b,1], Peter Fabian [a], Daniela Gurska [a], Zbynek Kozmik [a,*]

[a] Institute of Molecular Genetics, Videnska 1083, 142 20 Prague 4, Czech Republic
[b] Dipartimento di Scienze della Terra, dell' Ambiente e della Vita, Università di, Genova, Viale Benedetto XV, 5, 16132 Genoa, Italy

ARTICLE INFO

Article history:
Received 18 March 2013
Received in revised form
18 July 2013
Accepted 19 July 2013
Available online 7 August 2013

Keywords:
Bmp signaling
Axial patterning
Cell fate
Chordates
Evolution

ABSTRACT

In chordates, early separation of cell fate domains occurs prior to the final specification of ectoderm to neural and non-neural as well as mesoderm to dorsal and ventral during development. Maintaining such division with the establishment of an exact border between the domains is required for the formation of highly differentiated structures such as neural tube and notochord. We hypothesized that the key condition for efficient cell fate separation in a chordate embryo is the presence of a positive feedback loop for Bmp signaling within the gene regulatory network (GRN), underlying early axial patterning. Here, we therefore investigated the role of Bmp signaling in axial cell fate determination in amphioxus, the basal chordate possessing a centralized nervous system. Pharmacological inhibition of Bmp signaling induces dorsalization of amphioxus embryos and expansion of neural plate markers, which is consistent with an ancestral role of Bmp signaling in chordate axial patterning and neural plate formation. Furthermore, we provided evidence for the presence of the positive feedback loop within the Bmp signaling network of amphioxus. Using mRNA microinjections we found that, in contrast to vertebrate Vent genes, which promote the expression of Bmp4, amphioxus Vent1 is likely not responsible for activation of cephalochordate ortholog Bmp2/4. *Cis*-regulatory analysis of amphioxus Bmp2/4, Admp and Chordin promoters in medaka embryos revealed remarkable conservation of the gene regulatory information between vertebrates and basal chordates. Our data suggest that emergence of a positive feedback loop within the Bmp signaling network may represent a key molecular event in the evolutionary history of the chordate cell fate determination.

Introduction

The early separation of cell fate domains precedes the final decision of ectoderm to become neural and non-neural as well as mesoderm to become dorsal and ventral during chordate development. The subdivision of the ectoderm into neural and epidermal is also present in early embryos of Ecdysozoa and Lophotrochozoa that together with vertebrates belong to bilaterians possessing a centralized nervous system. Recent molecular data support the hypothesis of 19th century naturalist Geoffroy Saint-Hilaire, who suggested that dorsal vertebrate and ventral invertebrate central nervous systems are homologous (Arendt and Nubler-Jung, 1994; Denes et al., 2007). In *Drosophila*, the homolog of Bmp4 decapentaplegic (*dpp*) gene is required for patterning the dorsal non-neural ectoderm and the homolog of Chordin short gastrulation (*sog*) gene plays a crucial role in ventral neural cell fate specification (Biehs et al., 1996; Mizutani et al., 2006). Similarly to vertebrate homologs, Sog functions as a diffusible morphogen, which interacts with the Dpp signaling molecule, preventing it from binding to the receptors and from inducing strong Dpp/Bmp signaling. A high level of Bmp signaling is required to activate epidermal and repress neural genes in the ectoderm (Biehs et al., 1996; von Ohlen and Doe, 2000). One of the key conditions for inducing a high level of Bmp signaling is the ability to activate the expression of Dpp with the Dpp signaling itself (Biehs et al., 1996). Lower Bmp levels can diffuse into neural ectoderm and contribute to subdivision of the neural domains (Ashe et al., 2000; Mizutani and Bier, 2008; Mizutani et al., 2006). As in vertebrates and *Drosophila*, in early *Platynereis* embryo Bmp2/4 is devoid to be expressed from ventral neuroectoderm and neural patterning genes are sensitive to Bmp signaling (Denes et al., 2007). Although Bmp–Chordin network was shown to be present in hemichordates, it is not used to segregate epidermal and neural ectoderm (Lowe et al., 2006).

* Corresponding authors. Fax: +420 241063125.
E-mail addresses: kozmikova@img.cas.cz (I. Kozmikova), kozmik@img.cas.cz (Z. Kozmik).
[1] Equal contribution.

0012-1606/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ydbio.2013.07.021

In *Xenopus* and zebrafish, the molecular mechanisms of cell fate domain separation in the ectoderm resemble those described in *Drosophila*. Together with Bmps, members of the Vent gene family are key ventralizing factors in early *Xenopus* and zebrafish embryo (Dale et al., 1992; Henningfeld et al., 2000; Hoppler and Moon, 1998; Imai et al., 2001; Nguyen et al., 1998; Onichtchouk et al., 1996, 1998). It was shown that the expression of zebrafish Vent and Vox genes is initiated at mid-blastula stages by maternal gene Runx2 (Flores et al., 2008). The available data further suggest the existence of a Bmp4-positive feedback loop in *Xenopus* (Kim et al., 1998; Metz et al., 1998) and Bmp2/Bmp7 positive feedback loop in zebrafish (Kishimoto et al., 1997; Nguyen et al., 1998; Schmid et al., 2000). Additionally, Xvent-2, which is a homolog of zebrafish Vox, was shown to directly up-regulate the Bmp4 gene, thus playing an important role in the Bmp positive feedback loop (Schuler-Metz et al., 2000).

Recently, the Bmp–Chordin network was extensively studied in sea anemone *Nematostella* possessing a diffuse nervous system (Matus et al., 2006; Rentzsch et al., 2006; Saina et al., 2009). Interestingly, gain-of-function and loss-of-function experiments suggest inhibition of the expression of Dpp and Bmp5/8 as well as Chordin genes by Dpp in *Nematostella*. On the other side, Chordin of *Nematostella* prevents Bmp-like molecules from binding to its receptor, and, thus, activates the expression of Dpp through the double negative loop (Saina et al., 2009). To investigate the evolution of Bmp signaling pathway function during embryogenesis at the invertebrate-to-vertebrate transition, we therefore performed a study of the embryonic role of this pathway in amphioxus. Amphioxus is an amenable model organism to address questions regarding evolution of early development because of its basal position in the chordate tree. Moreover, amphioxus genomic, morphological, and developmental characteristics are probably highly similar to those of the chordate ancestor (Bertrand and Escriva, 2011). In amphioxus, the expression of the genes playing a key role in the dorsoventral patterning was shown to be conserved with the expression patterns of the homologous genes in *Xenopus* and zebrafish embryos (Yu et al., 2007). In particular, Bmp2/4 and Bmp5/8 are expressed in a mutually exclusive manner with Chordin in amphioxus (Yu et al., 2007). Using mammalian cell cultures we have recently shown in *in vitro* studies that the Bmp–Chordin gene regulatory network is conserved in amphioxus and that it was likely present in a common ancestor of all chordates (Kozmikova et al., 2011). Although gain-of-function experiments suggested the role of Bmp signaling in dorsoventral patterning of amphioxus (J.K. Yu et al., 2008; Yu et al., 2007), loss-of-function experiments have not yet been performed. The goal of this study was to find out whether Bmp signaling is required for axial patterning and for cell fate determination in the amphioxus embryo. In particular, we wanted to investigate whether the Bmp positive feedback loop is present in amphioxus similarly as in the case of vertebrates, where it is directly connected to early separation of cell fates leading to the development of a centralized nervous system and chorda.

Materials and methods

Amphioxus collection and embryo manipulation

Adults of Florida amphioxus (*Branchiostoma floridae*) were collected from Old Tampa Bay, Florida, during the summer breeding season. Adults were induced to spawn by electrostimulation as described (Yu and Holland, 2009a). Embryos were raised in the laboratory on site. All pharmacological treatments were performed on *B. floridae* unless stated otherwise. The embryos at early gastrula or early blastula stage were treated with Bmp inhibitor dorsomorphin (10 or 20 μM; Calbiochem), Bmp inhibitor LDN-193189 (3 μM, Axon Medchem), activin inhibitor SB505124 (30 μM; Sigma) or with 250 ng/μl of human recombinant BMP2 (R@D). Control embryos were administered to 10 or 20 μM of DMSO (Dimethyl Sulfoxide; Sigma). The control and treated embryos were fixed at larva stage for the analysis of the phenotypes and at early hatching neurula or late neurula stage for *in-situ* hybridization. The measuring of the length of the larva head region (the region anterior to pigment spot) or trunk (the region posterior to pigment spot) was done with Lucia image program. Statistical significance was determined using Student *t*-test in Microsoft Excel. Microinjection of eggs with mRNA was performed using European amphioxus (*Branchiostoma lanceolatum*). Adults of *B. lanceolatum* were collected in Banyuls-sur-mer, France, during the summer breeding season. The spawning of European amphioxus male and females was induced by shifting of the temperature as described (Fuentes et al., 2007).

RNA purification and real-time quantitative RT-PCR (qRT-PCR)

Developing embryos of *B. floridae* were collected into RNA later (Ambion). Embryos were treated with DMSO (control) or 20 μM dorsomorphin at blastula stage and allowed to develop until early neurula stage. Total RNA was extracted from approximately 250 embryos per experimental condition.

Standard procedures were used for RNA purification and reverse transcription. Briefly, total RNAs were isolated from embryos using the Trizol reagent (Invitrogen); contaminating genomic DNA was eliminated by DNAse I digestion and RNA was repurified using RNeasy Micro kit (Qiagen). Random-primed cDNA was prepared in a 20 μl reaction from 100 ng of total RNA using SuperScript VILO cDNA Synthesis kit (Invitrogen). cDNAs were produced from at least two independent RNA isolations and the PCR reactions were performed in triplicate for each primer set. Control reactions (containing corresponding aliquots from cDNA synthesis reactions that were performed without reverse transcriptase; minus RT controls) were run in parallel. PCR reactions were run using the LightCycler 480 Real-Time PCR System (Roche). Typically, a 10 μl reaction mixture contained 5 μl of LightCycler 480 SYBR Green I Master mix (Roche), 1 μl of primers (final concentration 0.5 μM) and cDNA diluted in 4 μl of deionized water. Crossing-threshold (CT) values were calculated by LightCycler® 480 Software (Roche) using the second-derivative maximum algorithm. The specificity of each PCR product was analyzed using the in-built melting curve analysis tool for each DNA product identified; additionally, PCR products were verified by sequencing. All primers were calculated using Primer 3 computer services at http://frodo.wi.mit.edu/. The housekeeping gene encoding ribosomal protein L32, *AmphiRPL32*, was used as internal control gene to standardize the quality of different cDNA preparations. The following genes were detected: Vent1, Vent2, Bmp2/4, Bra, Hex, and SoxB1a (see Supplementary Table 1 for sequences of primers used).

Whole-mount in situ hybridization in amphioxus embryos

Whole-mount *in situ* hybridization with digoxigenin-labeled RNA probes was done as described (Yu and Holland, 2009b). RNA probes were synthesized *in vitro* with T7/Sp6 RNA polymerase (Roche) using plasmid DNA that encoded the following amphioxus genes: Vent1, Vent2, Bmp2/4, Admp, Goosecoid, Chordin, Evx, Brachyury, Hex, Runx, Brn1/2/4, SoxB1a, SoxB1c. Plasmids were either previously described (Candiani et al., 2002; Kozmik et al., 2001; Meulemans and Bronner-Fraser, 2007; Yu et al., 2007) or were generated by cloning the corresponding cDNA from amphioxus cDNA library into pCR-BluntII-TOPO (Invitrogen) using primers summarized in Supplementary Table 1. To improve the

ISH protocol, after proteinase K and glycine treatments, embryos were transferred in special nests (millicell-PCF 12 µm, Millipore) and positioned in a 24-well plate for all subsequent steps. For color reaction the samples were transferred in a four-well plate. Color was developed by incubation in BM purple (Roche). For every treatment 12–15 embryos per gene were analyzed by *in situ* hybridization. Labeled whole mount embryos were photographed using an Olympus IX71 microscope.

Whole-mount in situ hybridization of medaka embryos

Medaka embryos of the Cab inbred strain were used for all experiments. Embryonic stages were determined according to Iwamatsu (2004). Embryos were fixed overnight in 4% PFA/PTW (PBS/Tween) at 4 °C and subsequently dechorionated manually. Whole-mount *in situ* hybridizations were performed at 65 °C using DIG-labeled RNA probes. The color reaction was carried out with NBT/BCIP, followed by re-fixation in 4% PFA/PTW. Embryos were cleared in 100% MeOH, rehydrated in PTW and mounted in 87% glycerol through a graded series of glycerol/PTW. The following probes were used: OlBmp4, OlChordin and OlAdmp (see Supplementary Table 1 for sequences of primers used to generate the corresponding plasmid).

Microinjection of mRNA into amphioxus and medaka eggs

RNA for injection was prepared using the mMessage mMachine kit (Ambion). The coding regions of medaka OlVent, amphioxus Vent1 (from *B. floridae*) and mCherry were subcloned into the pCS2+ expression vector to produce the plasmids pCS-OlVent, pCS-AmphiVent1 and pCS-mCherry, respectively. Microinjection of amphioxus eggs was generally done as described in Onai et al. (2010) with some modifications. For microinjection of amphioxus, a solution of 15% glycerol, 0.5% of Phenol Red (Sigma) plus 30 ng/µl of mRNA was used. For microinjection of medaka, a solution of 1xERM, 0.5% of Phenol Red and 20 ng/µl of mRNA were applied. Injection needles were pulled with a vertical pipette puller and back-filled with mRNA. In the case of amphioxus, the tips were broken slightly by touching the end of the needle with a pair of forceps. For medaka microinjection, the tips were broken by scratching the needle against the membrane of the embryo. A FemtoJet (Eppendorf) was used for pressure injection of RNA. Pressure was adjusted so that a bolus of RNA of about 2 nl volume was injected into one-cell stage medaka or 0.5 nl into amphioxus unfertilized eggs. Calibration was done with the Stage Micrometer. Injected embryos were fixed at the appropriate stage.

I-SceI-mediated transgenesis in medaka fish

I-SceI meganuclease transgenesis in medaka fish was performed as previously described (Rembold et al., 2006). Fertilized eggs of inbred Cab strain were collected immediately after spawning and were placed into cold (4 °C) 1 × Yamamoto's embryo rearing medium. Reporter genes designated AmphiChordin-GFP, AmphiAdmp-Cherry, AmphiBmp2/4-Long-Cherry and AmphiBmp2/4-Short-Cherry were generated by subcloning the putative 5' regulatory sequences of *B. floridae* genes into *I-SceI* transgenic vector (see Supplementary Table 1 for sequences of primers used to generate each plasmid). One-cell stage medaka embryos were injected with the solution containing *I-SceI*-based reporter gene plasmid and 0.25 U/µl *I-SceI* meganuclease in 0.5 × *I-SceI* buffer (New England Biolabs, USA)/0.5 × Yamamoto's embryo rearing medium. Final concentration of the injected plasmid was 10 ng/µl. The expression of the transgene was detected as early as at early gastrula stage (stage 13) in the case of AmphiChordin-GFP and at mid-gastrula stage (stage 16) in the case of AmphiAdmp-Cherry, AmphiBmp2/4-Long-Cherry, AmphiBmp2/4-Short-Cherry. The injecting setup was as follows: pressure injector Femtojet (Eppendorf); micromanipulator TransferMan NK (Eppendorf); borosilicate glass capillaries (GC100F10, Harward Apparatus); stereomicroscopes (Olympus SZX7, SZX9).

Immunohistochemistry

All incubation steps were carried out at room temperature. Specimens were transferred to 1 × PBS, 0.1% (vol/vol) Tween 20 (PBT) through 50% (vol/vol) and 25% (vol/vol) methanol in PBS. Specimens were washed three times (20 min each washing) in PBT, blocked in block solution [10% (wt/wt) BSA in PBS] for 1 h, and incubated with the antibody for Smad1/5 [pS463/pS465] (Catalog no. 700047, Invitrogen) in dilution 1:100 overnight at 4 °C. On the next day, specimens were washed three or four times in PBT (20 min each washing) and were incubated with secondary antibodies for 2 h. Secondary antibodies were washed away with three washings in PBT (20 min each washing). Nuclear counterstaining was carried out by incubation with 1 µg/ml DAPI in PBS and washing three times (5 min each washing). For fluorescence/confocal microscopy, the specimens were mounted in VECTASHIELD (Vector Laboratories, Inc.) using small coverslips as spacers between the coverslip and the slide. The confocal images were taken using a Leica SP5 confocal microscope and were deconvoluted and processed with SVI Huygens Image Deconvolution Software.

Whole mount detection of apoptosis on amphioxus embryos using TUNEL methodology

Detection of apoptotic DNA degradation was carried out with the ApopTag *in situ* Apoptosis Detection kit (Millipore, USA) as it was previously described in Bayascas et al. (2002).

Results

Bmp signaling activity during early embryogenesis of amphioxus

BMP activity gradients during development originate from complex interactions between BMP signaling components and its specific regulators. The mechanisms by which spatial regulation of BMP activity is achieved are varied in different organisms. To monitor the activity of Bmp signaling during early development of amphioxus, we performed immunostaining with the antibody which detects phosphorylated Smad1 or Smad5 (pSmad1/5). Bmp signaling induces specific phosphorylation of Smad1, Smad5 and Smad8. Interestingly, although a detectable level of mRNA expression of Bmp pathway ligands Bmp2/4, Bmp5/8 or Admp was first observed only at the onset of gastrulation in amphioxus embryo (Yu et al., 2007), we were able to detect pSmad1/5 already at two-cell stage (Fig. 1A–A″). At 32-cell stage the activity of Bmp signaling was seen ubiquitously (Fig. 1B–B″). A reduced signal was visible in some cells of the vegetal pole at 64-cell stage (Fig. 1C–C″, white arrows). At the late blastula stage we observed unequal distribution of the Bmp signaling activity within the embryo (Fig. 1D–D″). At mid-gastrula stage, the signal was absent in one region of the ectoderm adjacent to blastopore (Fig. 1E–E″), which is supposed to be the dorsal ectoderm, where the Bmp antagonists such as Chordin started to be expressed in the amphioxus gastrula embryo (Yu et al., 2007). It is interesting to note that although Bmp2/4 and Bmp5/8 are expressed mainly in the endoderm of the amphioxus mid-gastrula (Yu et al., 2007), pSmad1/5 is present in the ectoderm as well (Fig. 1E–E″). During early hatching neurula stage, a weak Bmp activity was observed in the most dorsal ectoderm and endomesoderm of the embryo (Fig. 1F–F″). At early mid-neurula stage, a strong pSmad1/5 signal was detected at the border regions of neural plate and in the ventral and posterior endoderm (Fig. 1G–G″).

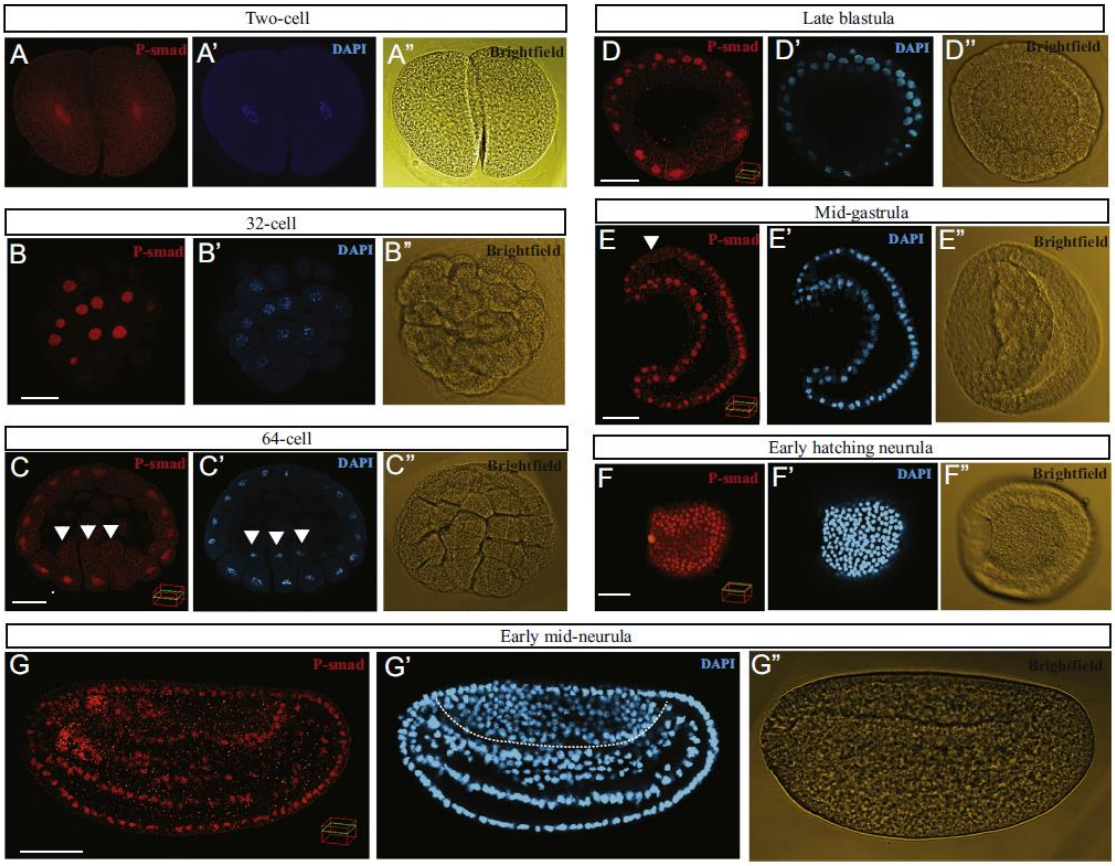


Fig. 1. pSmad1/5 immunostaining reveals localized BMP activity during embryogenesis of amphioxus. (A–A″) Two-cell stage. Z-stack of 32-cell (B–B″), 64-cell (C–C″), late blastula (D–D″), mid-gastrula (E–E″) and sagittal z-stack of early mid-neurula stage (G–G″). Three merged coronal z-stacks from the most dorsal side of early hatching neurula (F–F″). Green squares in the red boxes situated at the bottom right of C, D, E, F, and G show the approximate position of the z-stack. In B–C″ the animal pole is top and vegetal is bottom. Lateral view in (E–E″ and G–G″). Dorsal view in (F–F″). Anterior to the right in (F–F″ and G–G″). White arrowheads in (C′) demarcate the nuclear signal as compared to the absence of pSmad1/5 signal in (C) in the same cells. The arrowhead in (E) defines the probable position of dorsal ectoderm. The white line in (G′) demarcates the approximate border of the neural plate. Scale bar is 40 μM.

Our observations indicate that although the Bmp activity is initiated during gastrulation, pSmad1/5 has been detected from two-cell stage, suggesting the presence of Bmp ligands and receptors in amphioxus embryo before the onset of mid-blastula transition and gastrulation.

Inhibition of Bmp signaling induces the range phenotypes in amphioxus

Gain-of-function experiments suggested the role of Bmp signaling in the axial patterning of amphioxus (J.K. Yu et al., 2008; Yu et al., 2007); however, the loss-of-function experiments have not yet been performed. To interrogate the role of Bmp signaling in early development of amphioxus, we treated the embryos at early blastula or early gastrula stage with 10 μM or 20 μM doses of dorsomorphin (DM). DM is a small molecule inhibitor of Bmp signaling which selectively prevents phosphorylation of Bmp-responsive Smad1/5/8 (P.B. Yu et al., 2008). To implement the gain-of-function experiment, the activation of Bmp signaling, we treated animals with purified recombinant human BMP2 protein. The extent of severity varied depending on the dose and timing (Fig. 2A–D, G). The embryos exposed to higher doses of DM (20 μM) at early blastula stage showed the most severe phenotype as compared to controls (compare Fig. 2A and C, D, G). While the head looked normal, the trunk was shorter and the tail did not elongate properly (Fig. 2G). In contrast, the treatment with BMP2 led to severe headless phenotypes (Fig. 2E and F), corroborating previously published results (Yu et al., 2007). In the early amphioxus larvae, which were treated with 20 μM of DM at early blastula stage, we observed the upregulation of the axial mesoderm marker Brachyury in the both posterior and anterior domains (Supplementary Figs. S1A and S1B). The expression of neural tube and anterior endoderm marker SoxB1c was expanded laterally and posteriorly (Supplementary Figs. S1C and S1D). To directly compare the effect of inhibition of Bmp signaling in amphioxus and vertebrates, the embryos of teleost fish medaka were treated with the same concentrations of DM at early blastula stage. We observed a dorsalized phenotype, comparable to those previously described for zebrafish (P.B. Yu et al., 2008). Like in the case of amphioxus, medaka embryos displayed defects of the tail and a shorter trunk (Supplementary Figs. S1E and S1F).

Bmp signaling is required for the correct expression pattern of early ventral and dorsal markers in amphioxus early neurula

High level of Bmp signaling activity was detected at early hatching neurula stage (Fig. 3A–A″). To further investigate the

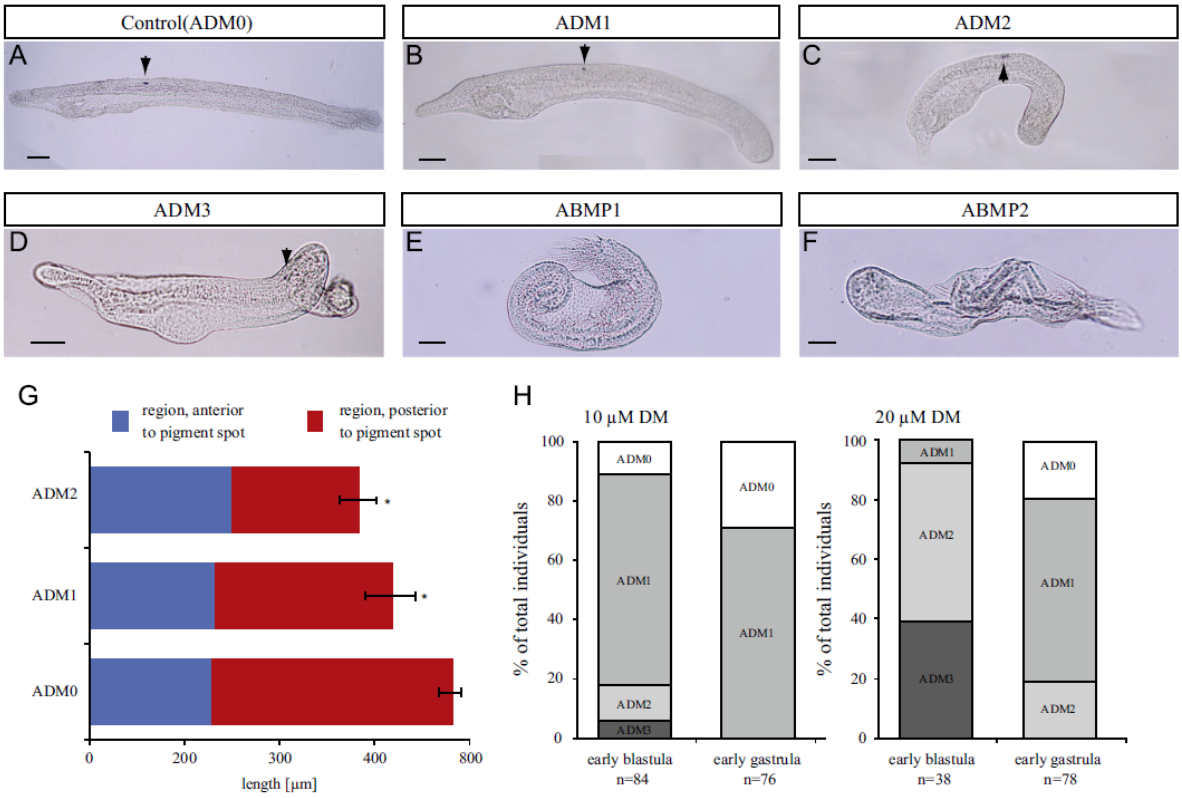


Fig. 2. The perturbation in Bmp signaling causes the range of phenotypes in amphioxus embryos. Two-days-old amphioxus larvae that were treated with 10 μM or 20 μM DM at early blastula or early gastrula stage show weak (B, C) or strong (D) phenotypes as compared to the control larvae (A). DM treatment leads to the development of the larvae with shortened trunk region with a minimal effect on the length of the head region (G, $^{*}P < 0.05$, $n = 10$ in each group). The length of the trunk region measured from the first pigment spot was considered as the criteria for ADM1 (B) and ADM2 (C) phenotypes. The embryos with ADM3 (D) phenotype had the shortest trunk as compared to the head measured from the first pigment spot. The percent of individuals with defined phenotypes is summarized in (H). The arrows denote the first pigment spot in (A–D). BMP2 treatment at the early blastula stage resulted in headless amphioxus larvae with either shortened and curved tails (phenotype ABMP1 (E)) or shortened and straight tails (phenotype ABMP2 (F)).

effect of DM on amphioxus axial patterning we examined the expression pattern of dorsal and ventral markers at early neurula stage in the embryos treated with 20 μM of DM. First, we detected decrease of the pSmad1/5 signal in the embryos treated with 20 μM of DM that demonstrated downregulation of Bmp signaling in our experiments (Fig. 3B–B″). In contrast, treatment of the embryos with BMP-2 led to the ubiquitous activity of Bmp signaling (Fig. 3C–C″) as compared to controls (Fig. 3A–A″). Our results showing effective manipulation of Bmp signaling pathway in amphioxus embryos are consistent with recently published data (Lu et al., 2012). Dorsal markers Chordin and Goosecoid were conspicuously upregulated in DM treated embryos (Fig. 3D, D' and E, E'). Chordin was expanded laterally and spread throughout the embryo (Fig. 3D'). Goosecoid was partially expanded in dorsal mesoderm and highly upregulated in anterior dorsal mesoderm (Fig. 3E'). Conversely, exposure to BMP2 resulted in the elimination of Chordin mRNA expression in the whole embryo and downregulation of Goosecoid (Fig. 3D" and E"). Exposure to 20 μM DM also led to the decrease of Evx expression (Fig. 3F') The treatment with BMP2 resulted in the upregulation of Evx (Fig. 3F"), as compared to the control (Fig. 3F). We next examined the expression of Vent1 and Vent2 genes that were shown to be direct targets of Bmp signaling (Kozmikova et al., 2011) in control and pharmacologically treated embryos. Vent1 and Vent2 genes in amphioxus originated from lineage-specific duplication and show 83% nucleotide identity within their entire open reading frames. Because of the high degree of sequence identity, previous studies were not able to distinguish the expression pattern of the two genes by whole-mount *in situ* hybridization with an RNA probe derived from the coding region of Vent1 (Kozmik et al., 2001; Yu et al., 2007). In order to distinguish the expression of Vent1 and Vent2 we synthesized probes from their divergent 3' regions of mRNA. Although the expression of amphioxus Vent genes overlaps quite broadly in lateral domains of the embryo demarcating the neural plate, the expression of Vent1 is stronger in the posterior domain at early neurula stage (Fig. 3F and G). Vent1 as well as Vent2 were downregulated and an altered spatial pattern of both genes was detected in DM-treated embryos (Fig. 3G' and H'; Supplementary Fig. S2). Similarly to Evx (Fig. 3F") BMP2 treatment caused upregulation of expression of Vent1 and Vent2 genes (Fig. 3G" and H"). Since the data showing an important role of Tbx2/3 in mediating Bmp signaling during early development of sea urchin have recently been published (Lapraz et al., 2009), we next decided to investigate the amphioxus Tbx2/3 expression in embryos perturbed in Bmp signaling. The expression of amphioxus Tbx2/3 at the early neurula stage is present in ventral posterior endoderm, posterior ectoderm, and paraxial mesoderm and in the posterior neural ectoderm (Fig. 3I). Both the inhibition and the activation of Bmp signaling resulted in specific alterations in the normal spatial expression pattern of amphioxus Tbx2/3. DM

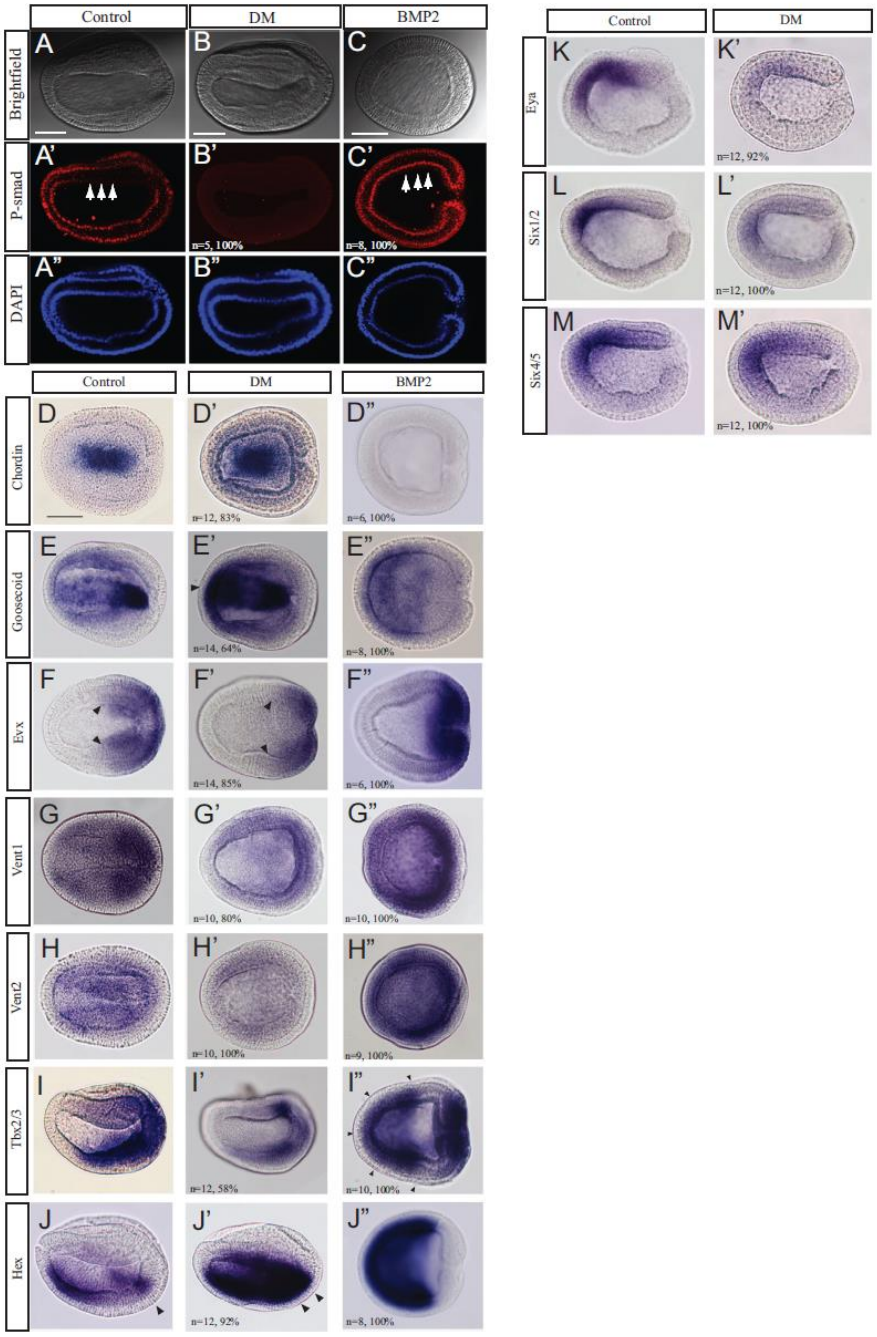


Fig. 3. Dorsal and ventral marker genes are affected in the embryos at early hatching neurula treated with 20 μM DM or 250 ng/ml of BMP2 at early blastula stage. pSmad1/5 immunostaining of control (A–A″) and DM-treated embryos (B–B″) showing effective downregulation of Bmp signaling. In contrast, the treatment with BMP2 resulted in the ubiquitous activity of BMP signaling (C–C″). Chordin expression is expanded in DM-treated embryos (D′) and downregulated in BMP2-treated animals (D″) as compared to controls (D). Goosecoid, which is expressed in dorsal mesoderm of control embryos (E), is upregulated in the DM-administered embryos (E′) and remains weakly expressed in the anterior domain of BMP2-treated embryos (E″). DM treatment leads to the downregulation of Evx (F′) in the regions, bordering dorsal mesoderm and neural plate, as compared to control embryos (F). Treatment with DM resulted in the downregulation of Bmp signaling target genes Vent1 (G′), Vent2 (H′) and Tbx2/3 (I′) as compared to controls (G, H and I, respectively). Bmp treatment led to the upregulation of expression of Evx (F′), Vent1 (G″), Vent2 (H″) and Tbx2/3 (I″). The Hex gene expressed throughout endomesoderm in control embryos (J) is upregulated in the posterior endoderm of DM-treated animals (J′) and in anterior endoderm of BMP2-treated embryos (J″). The treatment with DM resulted in the downregulation of amphioxus Eya (K′), Six1/2 (L′) but not Six4/5 gene as compared to controls (K, L, M). The embryos in (A–G″, H″, I″), are shown in a dorsal view and in (H and H′, I and I′, J–M′) in a lateral view. The arrowheads demarcate the expression changes described.

treatment led to the inhibition of Tbx2/3 expression mainly in the neural ectoderm and paraxial mesoderm (Fig. 3I'). The activation of Bmp signaling resulted in an essential alteration in the spatial expression of amphioxus Tbx2/3 that was manifested by the emergence of an ectopic anterior expression domain (Fig. 3I''). These data suggest that, similarly as in the case of sea urchin, Bmp signaling is required to maintain amphioxus Tbx2/3 expression during early development. We further investigated whether the expression of ventral anterior endoderm marker Hex was changed in the embryos administered to 20 μM DM at early blastula stage. Inhibition of Bmp signaling led to upregulation of the Hex gene (compare Fig. 3J and J'; Supplementary Fig. S2). It appeared that the ventral domain of Hex expression was enhanced more than the anterior domain in DM-treated animals (Fig. 3J'). These results suggest that negative regulation of Hex gene by Bmp signaling that occurs during the gastrulation in *Xenopus* embryo (Rankin et al., 2011) may take place in amphioxus as well. We observed a broad expression of Hex genes in the anterior region of BMP2-treated embryos (Fig. 3J''). The same expression pattern of the Hex gene has recently been shown in the embryos injected with mRNA of amphioxus Bmp2/4 (Onai et al., 2010). These data suggest that ectopic activation of Bmp signaling might promote differentiation of the anterior domain of the embryo into anterior endoderm.

To verify the specificity of DM we treated the embryos at an early blastula stage with another inhibitor of Bmp signaling pathway LDN-193189. Control and treated embryos were analyzed at early hatching neurula stage by whole mount *in situ* hybridization. As in the case of the treatment with DM, we observed the upregulation of amphioxus Chordin, Goosecoid and Hex genes (Supplementary Figs. S3A', S3B' and S3E') as well as downregulation of Vent1 and Evx (Supplementary Figs. S3C' and S3D') in the embryos treated with LDN-193189 as compared to controls (Supplementary Figs. S3A, S3B, S3E, S3C and S3D, respectively). To check whether the cell death is not responsible for the changes in the gene expression patterns in the embryos treated with 20 μM of DM, we applied the terminal deoxynucleotidyl transferase (TdT) mediated dUTP nick end-labeling (TUNEL) technique. The TUNEL-positive cells were not detected in the embryos treated with DMSO or DM (Supplementary Figs. S4A and S4B).

Finally, we investigated the role of Bmp signaling in the regulation of amphioxus Eya, Six1/2, Six4/5, which are orthologues of vertebrate preplacodal markers Eya, Six1 and Six4. The early requirement of Bmp signaling for the expression of preplacodal markers and the specification of preplacodal domain within the non-neural ectoderm in zebrafish is well established (Kwon et al., 2010). In contrast to vertebrate preplacodal markers, expressed in the anterior non-neural ectoderm by late gastrulation, amphioxus Six1/2, Six4/5 and Eya are expressed in the anterior endoderm (Fig. 3K, L and M) (Kozmik et al., 2007). It was demonstrated that Bmp signaling activity before gastrulation is required for the expression of vertebrate Eya1 and Six4 (Kwon et al., 2010), establishing preplacodal competence throughout the non-neural ectoderm. In amphioxus, DM treatment led to significant

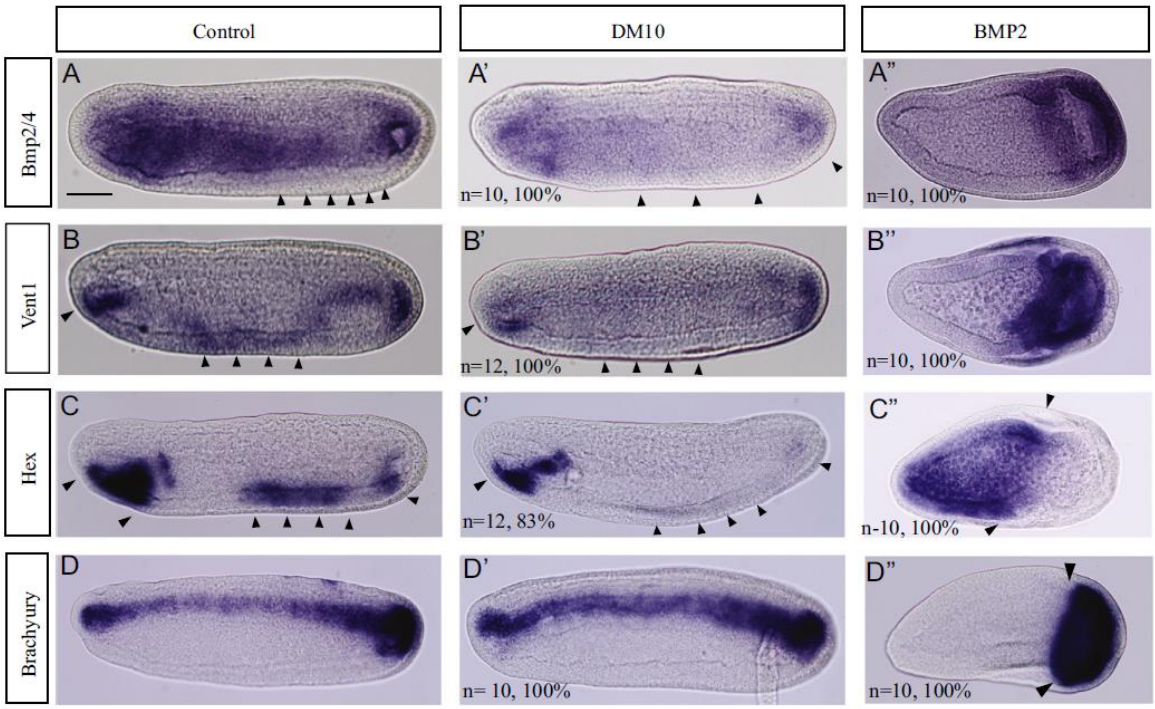


Fig. 4. Treatment with low doses of DM (10 μM) at early blastula stage results in the change of cell fate in the posterior and ventral mesoderm of late amphioxus neurula. The expression of amphioxus Bmp2/4 (A), Vent1 (B), Hex (C) and Brachyury (D) at late neurula stage in the control embryos. Panels A′, B′, C′, D′ represent embryos at late neurula stage treated with 10 μM DM from the early blastula stage. Panels A′′, B′′, C′′, D′′ show embryos administered 250 ng/ml of BMP2 at early blastula stage. Inhibition of Bmp signaling caused downregulation of Bmp2/4 at late neurula stage with stronger effect in the ventroposterior domain of endoderm and mesoderm (A′). Conversely, ectopic activation of Bmp signaling led to the upregulation of Bmp2/4 expression in the posterior domain of the embryo (A′′). The marker of ventral mesoderm Vent1 (B′) and endothelial marker Hex (C′) disappeared from ventral domains and were reduced in anterior endoderm of DM-treated embryos. BMP2 treatment resulted in strong upregulation of Vent1 in the posterior endomesoderm (B′′) and Hex in the anterior endoderm (C′′). The expression of notochord marker Brachyury was not significantly changed in the embryos treated with 10 μM DM (D′). BMP2 treatment caused strong upregulation of Brachyury in the posterior mesoderm and complete loss from the dorsal mesoderm (D′′). All embryos are shown from lateral views. The arrowheads demarcate the expression changes described. Scale bar is 50 μM.

downregulation of Eya (Fig. 3K') and Six1/2 (Fig. 3L'), but not Six4/5 (Fig. 3M'). This discrepancy of regulation of Six/Eya in amphioxus and vertebrates may reflect a derived process in cephalochordates, rather than a vertebrate invention. Combined, these data suggest that Bmp signaling is required for correct cell fate specification and axial patterning in the amphioxus.

Bmp signaling is required for ventral and posterior cell fate specification at the late neurula of amphioxus

Although the treatment with 10 μM DM at early blastula stage manifested in obvious phenotype of amphioxus two days-old larva (Fig. 2), we did not observe significant changes in the expression of ventral and dorsal markers at early neurula stage in these embryos (data not shown). Therefore we further investigated whether the inhibition of Bmp signaling with low dose (10 μM of DM) affects cell fate determination at the late neurula stage. Indeed, conspicuous alterations of the expression pattern of several genes were observed in these embryos at the late neurula stage. It is interesting that low-dose treatment (10 μM DM) resulted in a significant reduction of Bmp2/4 expression in the ventral and posterior domains of the embryo at late neurula stage (compare Fig. 4A and A'). Further, we observed that Vent1 expression was greatly suppressed in ventral mesoderm and downregulated in anterior endoderm (Fig. 4B and B'). Amphioxus Hex gene, which is expressed in anterior and ventroposterior endoderm in the control embryos (Fig. 4C), was lost in ventroposterior and reduced in anterior domains (Fig. 4C'). The treatment with BMP2 led to upregulation of the posterior expression domain of Bmp2/4 (Fig. 4A'') and Vent1 (Fig. 4B'') and to upregulation of the anterior expression domain of the Hex gene (Fig. 4C''). These results indicate that although Bmp signaling is required for proper differentiation of ventral, posterior as well as anterior endoderm, ventral and posterior endodermal fate is more sensitive to the Bmp signaling knockdown. The expression of notochord marker Brachyury was not significantly changed in the embryos treated with the low dose of DM (Fig. 4D') as compared to the control embryos (Fig. 4D). In contrast, BMP2 treatment resulted in the upregulation of Brachyury in the posterior mesoderm (Fig. 4D''). In summary, our results indicate the requirement of Bmp signaling for the cell fate determination mainly in ventral and posterior domains of amphioxus late neurula. In addition, our data suggest that the use of low DM doses, likely causing only partial inhibition of Bmp signaling, allows bypass of the early "strong" phenotypes.

Inhibition of Bmp signaling leads to the expansion of neural plate markers in amphioxus

In vertebrates, downregulation of Bmp signaling is absolutely required for the establishment of neural fate of embryonic cells (Marchal et al., 2009). Similarly to vertebrates, the dorsal ectoderm in amphioxus differentiates into neural plate during gastrulation. To test whether more effective inhibition of Bmp signaling (using 20 μM DM) may cause alterations of the neural cell fate establishment in amphioxus we examined the expression of early neural plate markers, Brn1/2/4, SoxB1a and SoxB1c genes (Candiani et al., 2002; Holland et al., 2000; Meulemans and Bronner-Fraser, 2007) at early hatching neurula stage. Treatment with 20 μM DM at early blastula stage resulted in the expansion of the SoxB1a gene into the dorsal ectoderm and in its ectopic expression in non-neural ectoderm throughout the embryo (compare Fig. 5A and B; Supplementary Fig. S2). In the case of Brn1/2/4 we did not detect strong local upregulation following DM treatment, but instead, the expression of the Brn1/2/4 gene expanded within the embryo without apparent establishment of the neural plate domain (compare Fig. 5C and D). We also observed upregulation of SoxB1c manifested by the expansion of the signal and by the elimination of the gap between anterior and posterior expression domains (compare Fig. 5E and F) at early neurula stage. In contrast, the marker of non-neural ectoderm AP-2 was downregulated in the embryos treated with the inhibitor of Bmp signaling (compare Fig. 5G and H). Taken together, our results indicate that inhibition of Bmp signaling results in the expansion of neural fate and the constriction of epidermal fate in the ectoderm of amphioxus embryo, and are thus consistent with an ancestral role of Bmp signaling in the chordate neural cell fate determination.

Positive feedback loop within the Bmp signaling GRN functions during gastrulation in medaka and amphioxus

Quadruple knockdown of Bmp2, Bmp4, Bmp7 and Admp in *Xenopus* resulted in ubiquitous central nervous system formation (Reversade and De Robertis, 2005). Our analysis of neural plate specific marker genes in DM-treated embryos indicated that in amphioxus, similarly as in vertebrates, the inhibition of Bmp signaling leads to the expansion of neural plate borders. In zebrafish, mutations in Bmp2b and Bmp7 genes are manifested by strong dorsalized mutant phenotypes and lead to the downregulation of Bmp2b and Bmp4 in ventral and lateral regions (Nguyen et al., 1998), suggesting that in vertebrates Bmp autoactivation is required for maintaining the ventral and lateral non-neural cell fate during gastrulation. The data from *Xenopus* and zebrafish allowed us to hypothesize that treatment of vertebrate medaka embryo with DM should result in the reduction of Bmp4 expression. This would (i) provide more evidence for a universal use of Bmp positive feedback loop among vertebrates, and (ii) allow a direct side by side comparative analysis of the underlying GRN between amphioxus and medaka. Indeed, treatment of medaka embryo with 20 μM DM at early blastula stage led to strong suppression of the Bmp4 level in medaka embryo at mid-gastrula stage (Fig. 6A and B). Another member of the Bmp family, Admp, known to have ventralizing activity in *Xenopus* (Dosch and Niehrs, 2000; Moos Jr. et al., 1995) and zebrafish (Lele et al., 2001; Willot et al., 2002), despite being expressed in the dorsal gastrula organizing center, was expanded in DM-treated embryos as compared to controls (Fig. 6C and D). In contrast, Chordin was strongly upregulated and expanded in lateral and ventral domains of embryos exposed to 20 μM DM (Fig. 6E and F). We then investigated whether the Bmp signaling positive feedback loop is also part of GRN in the basal chordate amphioxus. Analysis of Bmp2/4 and Bmp5/8 expression in embryos treated with 20 μM DM from early blastula stage revealed that indeed, Bmp signaling positive feedback loop functions during amphioxus gastrulation. The expression of Bmp2/4 was downregulated in embryos inhibited in Bmp signaling (Fig. 7A and B; Supplementary Fig. S2). In contrast to medaka Bmp4, however, the expression of amphioxus Bmp2/4 was almost not affected in endomesoderm but instead was most conspicuously reduced in the ectoderm that is best seen using blastopore view of the embryos (Fig. 7C and D). The Bmp5/8 gene, expressed in the endomesoderm of control embryos (Fig. 7E), was strongly downregulated in the emryos administrated to DM (Fig. 7F). We next examined the expression of amphioxus Admp that is expressed dorsally (Fig. 7G), similarly as its vertebrate orthologs, thus matching the pattern of organizer-specific genes Chordin and Goosecoid during gastrulation (Yu et al., 2007). Inhibition of Bmp signaling by DM resulted in slight expansion of an anterior Admp expression domain; significantly, no expansion of a dorsal domain of Admp was detected (Fig. 7H). Since it was recently shown that Nodal and Bmp signaling have the opposite role in axial patterning of an amphioxus embryo (Onai et al., 2010), we were interested to see whether inhibition of Nodal signaling leads to a change in the expression of Bmp2/4 and Admp. When

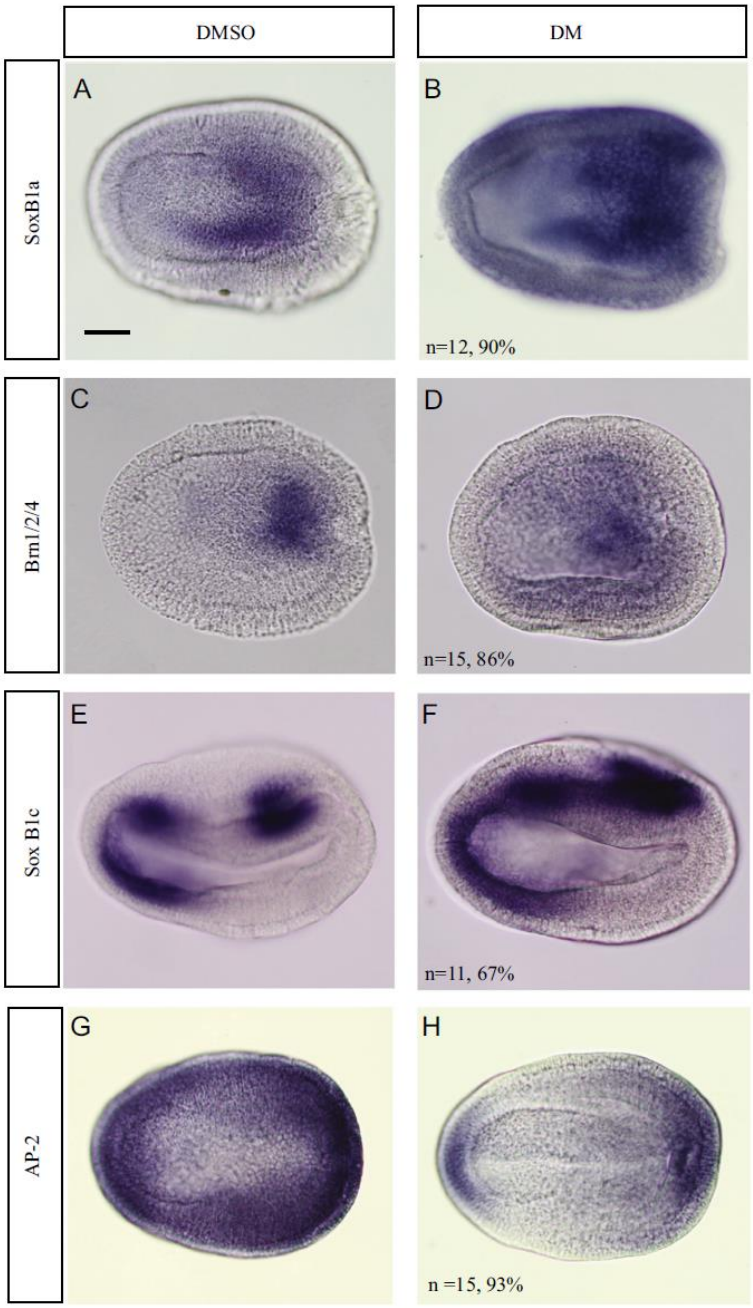


Fig. 5. Inhibition of Bmp signaling affects neural plate markers. Treatment with 20 μM DM at early blastula stage was manifested by expanded expression of SoxB1a (B), Brn1/2/4 (D) SoxB1c and (F) as compared to control embryos (A, C, E). In contrast, exposure of embryos to 20 μM DM at early blastula stage resulted in downregulation of the non-neural ectoderm marker AP-2 (H) as compared to control (G). The embryos in (A–D, G–H) are shown in a dorsal view and in (E) and (F) in a lateral view. Scale bar is 50 μM.

Nodal signaling was inhibited at early blastula stage using chemical inhibitor SB505124, widespread ubiquitous expression of amphioxus Bmp2/4 as well as Bmp signaling target genes Vent1 and Vent2 was detected (compare Fig. 7I and J, M and N and O and P). These results indicate that Nodal signaling is not required for initiation of Bmp2/4 expression like it was shown in the case of sea urchin (Lapraz et al., 2009) but instead, it is necessary for correct spatial patterning of the Bmp2/4 in amphioxus. In contrast, we observed complete downregulation of dorsally expressed Admp in the embryos treated with the inhibitor of Nodal signaling (compare Fig. 7K and L). Combined, our data provide evidence for the presence of a positive feedback loop within the Bmp signaling GRN

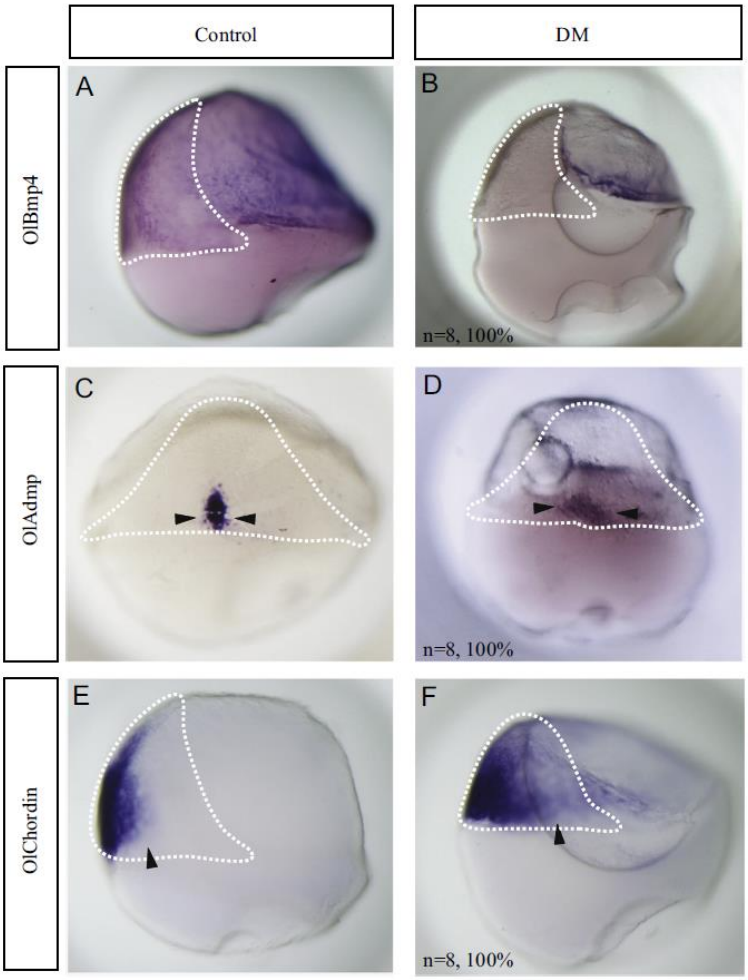


Fig. 6. Positive feedback loop in the Bmp signaling GRN functions during gastrulation in medaka. Bmp4 expression was greatly suppressed at mid-gastrula stage in the embryos treated with 20 μM DM at early blastula stage (A and B). Treatment with 20 μM DM resulted in slight widening of the dorsal posterior expression domain of Admp (D) in comparison with the expression in control embryos (C). Conversely, administration of 20 μM DM led to strong upregulation of organizer marker Chordin, which was spread into lateral domains in the treated embryos (F) as compared to controls (E). White dashed line indicates position of the shield in medaka embryo. The arrowheads demarcate the expression changes described.

in medaka and amphioxus, which is indicative of an ancestral character of this mechanism in chordates.

The function of Vent genes in the Bmp autoactivation loop is not conserved in chordates

Previous data indicated that, in vertebrates, Vent genes, which are targets of Bmp signaling, are able to promote the expression of Bmp4, thus enhancing the Bmp autoregulatory feedback loop (Imai et al., 2001; Schuler-Metz et al., 2000). To study the functional role of Vent genes as mediators of Bmp autoactivation loop we performed overexpression studies in medaka and amphioxus. Injection of *in vitro* produced mCherry mRNA into eggs of both species demonstrated efficient translation in developing embryos and no effect on normal development (Fig. 8A and B). To confirm that Vent-mediated activation of Bmp signaling is indeed a general mechanism among vertebrates we injected medaka one-cell stage embryo with medaka Vent mRNA. Consistent with the data from other vertebrates, medaka Vent was able to promote the expression of medaka Bmp4 at gastrula stage. Specifically, Bmp4 expression was upregulated and spread into the dorsal shield of medaka embryos (compare Fig. 8C and C'). In contrast, Chordin expression was reduced and restricted to the most dorsal domain (compare Fig. 8D and D'). To investigate whether a similar GRN architecture, which ensures a lock-down mechanism for ventrolateral cell fate specification, is present in amphioxus, we injected amphioxus eggs with amphioxus Vent1 mRNA. It is interesting to note that microinjection of amphioxus eggs with mRNA of the Vent1 gene, shown previously to be a direct target of Bmp signaling (Kozmikova et al., 2011), manifested in a ventralizing phenotype quite reminiscent of the phenotype observed in embryos treated with BMP2 protein. In particular, at the early neurula stage the Vent1-injected embryos were round (Fig. 8E' and F') like the embryos that have undergone BMP2 treatment (Fig. 3C"–J"). At mid-neurula stage, Vent1-injected embryos had a narrowed anterior domain as compared to the enlarged posterior domain (Fig. 8G' and H'). This phenotype is further consistent with the observation of headless amphioxus

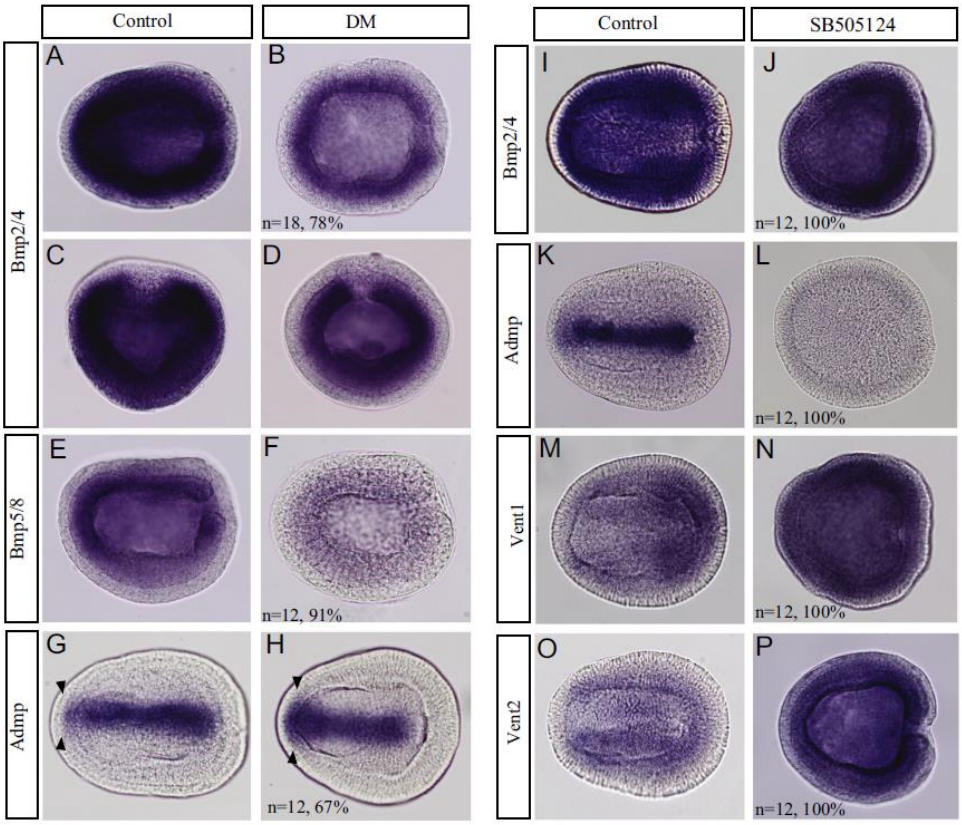


Fig. 7. Positive feedback loop within the Bmp signaling GRN is present in basal chordate amphioxus. The expression of amphioxus Bmp2/4 (A and C), Bmp5/8 (E), Admp (G), in the control embryos treated with DMSO. (B), (D), (F) and (H) represent the embryos at early neurula stage that were treated with 20 μM DM at early blastula stage. Application of 20 μM DM led to the conspicuous loss of Bmp2/4 expression within the ectoderm (B and D). Treatment with DM resulted in the overall downregulation of amphioxus Bmp5/8 gene (F). Slight widening of the anterior expression domain of Admp was observed in DM-treated embryos (compare position of arrowheads in G and H). Bmp2/4 (J), Admp (L), Vent1 (N) and Vent2 (P) expression in the embryos treated with the inhibitor of Nodal signaling SB431542 as compared to controls (I, K, M and O, respectively). The embryos in (A, B, G, H) are shown in dorsal views, (C) and (D) in blastopore view, (E) and (F) in lateral view.

larvae, which developed as a result of treatment with BMP2 at early blastula stage (Fig. 2E and F). However, in contrast to experiments in medaka (this study) and other vertebrates (Imai et al., 2001; Schuler-Metz et al., 2000), injection of amphioxus Vent1 mRNA into amphioxus egg did not promote dorsal expansion of amphioxus Bmp2/4 but rather slightly repressed its expression at early neurula stage (Fig. 8E and E'). The expression of Chordin was reduced dorsally (Fig. 8F and F'), consistent with our previously published data showing that amphioxus Chordin is a direct target of Vent1 transcriptional repressor (Kozmikova et al., 2011). We further investigated the expression of Bmp2/4 and Chordin in Vent1-injected embryos at the late mid-neurula stage. As compared to controls, Bmp2/4 expression was only slightly reduced in Vent1-injected embryos (compare Fig. 8G and G'), whereas an almost complete loss of Chordin expression was observed (compare Fig. 8H and H'). These data further support the conclusion that Vent1 does not activate Bmp2/4 in the basal chordate amphioxus. To test whether the divergence in Vent activity is due to a change in trans (evolution of Vent protein activity) or due to the evolution of *cis*-regulatory sequences for Vent targets, we analyzed the effect of overexpression of medaka Vent and amphioxus Vent1 genes in the context of medaka embryo. The microinjection of amphioxus Vent1 mRNA into medaka embryos led to the downregulation of medaka Chordin expression (Fig. 8D") like in the case of the injection of medaka Vent (Fig. 8D'). However, endogenous medaka Bmp4 was not upregulated in the shield of medaka embryos injected with amphioxus Vent1 mRNA (Fig. 8C") as compared to the embryos injected with medaka Vent mRNA (Fig. 8C'). These results suggest that the divergence in the Vent activity is caused by evolution of Vent protein. It is important to note that amphioxus Vent proteins do not show any significant sequence conservation with zebrafish, medaka and *Xenopus* Vent proteins outside of the homeodomain (data not shown). We have previously shown that, in contrast to Xvent2 (Gao et al., 2007), amphioxus Vent1 was not able to stimulate the Lef/Tcf-responsive reporter gene in the same transcriptional assay (Kozmikova et al., 2011), indicating that Vent1 does not possess a strong transactivation domain like its vertebrate orthologs. In summary, our data suggest that although the positive feedback loop in the Bmp signaling GRN is evolutionarily conserved among chordates, it is mediated by distinct molecular mechanisms in basal chordates and vertebrates.

Regulatory sequences of amphioxus Chordin, Admp and Bmp2/4 genes in the context of vertebrate embryo

Reports of successful transgenesis in amphioxus are scarce (Beaster-Jones et al., 2007; Holland et al., 2008; Yu et al., 2004).

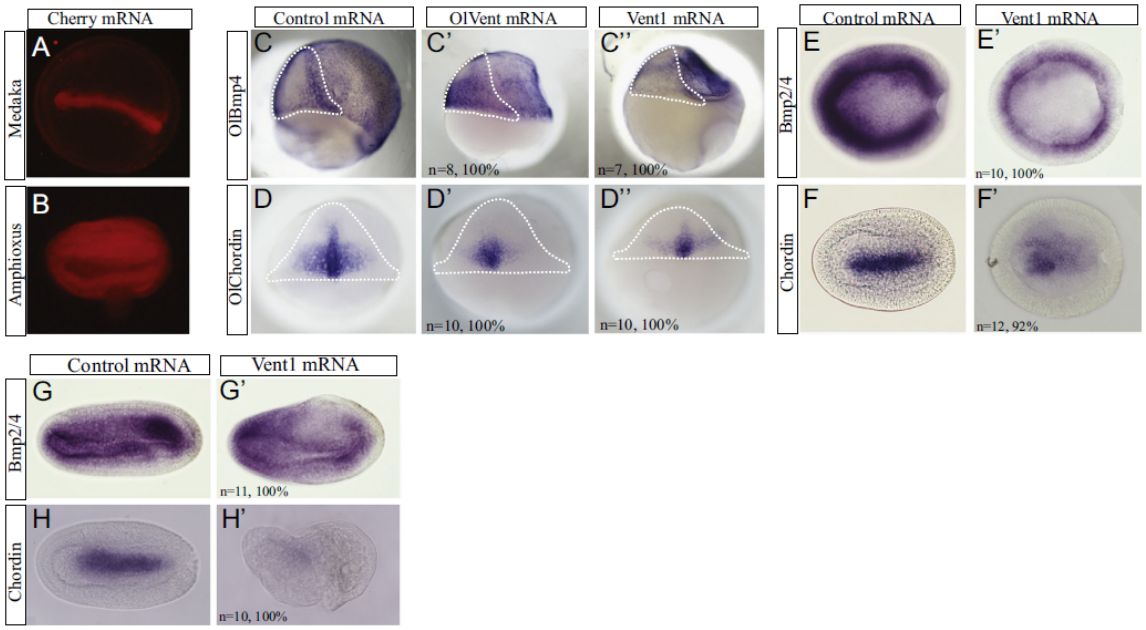


Fig. 8. The function of Vent genes in the Bmp autoactivation loop is not conserved in chordates. Injection of *in vitro* produced Cherry mRNA (control mRNA) into eggs of both medaka and amphioxus demonstrated efficient translation in developing embryos and had no effect on normal development (A and B). Microinjection of medaka Vent mRNA caused ectopic dorsal expansion of Bmp4 (C') but repression of medaka Chordin (D') at mid-gastrula stage as compared to the microinjection of control mRNA shown in (C) and (D), respectively. Microinjection of amphioxus Vent1 mRNA into medaka embryos did not lead to the expansion of medaka Bmp4 expression into the shield (C″) but resulted in the strong suppression of medaka Chordin expression (D″). Amphioxus Vent1 mRNA injected into amphioxus eggs caused the downregulation of both Bmp2/4 (E′) and Chordin (F′) expression in the embryos at early neurula stage as compared to the embryos injected with control mRNA (E and F, respectively). Microinjection of Vent1 led to slight suppression of Bmp2/4 (G, G′) and strong repression of amphioxus Chordin (H′) expression at mid-neurula stage as compared to the embryos injected with control mRNA (G and H). Lateral (C, C′, C″) and shield views (D, D′, D″) of medaka embryos are shown. Dorsal (E–F′, H and H') and lateral (G and G′) views of amphioxus embryos are shown. White dashed line indicates position of the shield in medaka embryo.

This is largely due to the fact that the current method of amphioxus transgenesis is inefficient and the few transgenic embryos obtained display a fairly high degree of mosaicism, making interpretation of data less reliable. We have previously shown that the promoters of amphioxus Vent1 and Vent2 genes are correctly regulated in medaka embryo and that their spatio-temporal activity is by large similar to the activity of orthologous vertebrate promoter (derived from the Xvent2b gene) in the same context (Kozmikova et al., 2011). Since we were not able to obtain reliable reporter gene expression in amphioxus using the previously characterized amphioxus Vent1 gene promoter (data not shown), we decided to employ the reliable and efficient medaka transgenesis in order to analyze putative regulatory sequences of additional amphioxus genes, comprising the Bmp–Chordin pathway. Furthermore, Yu et al. described high conservation of gene expression patterns between amphioxus and vertebrates during early development, suggesting the presence of an organizer in chordates (Yu et al., 2007), allowing us to make precise predictions about the areas of homology in the early embryos of amphioxus as compared to medaka. In this work we investigated whether the 5'regulatory region (putative promoter) of amphioxus Chordin gene, which is an early marker of vertebrate organizer, can be activated in the shield of medaka. In addition, we also analyzed the activity of putative promoters derived from amphioxus Admp and Bmp2/4 genes. Corresponding EGFP or mCherry reporter-based constructs (Fig. 9A–D) were injected into one-cell stage medaka embryos and their spatio-temporal expression was monitored during early medaka embryogenesis. Remarkably, the amphioxus Chordin promoter was activated in the shield (Fig. 9E and E'), precisely recapitulating the expression of endogenous chordin during gastrulation (Fig. 9E"). The reporter gene driven by Admp promoter was also active in the shield (Fig. 9F and F'), where the endogenous medaka Admp gene is expressed. However, in contrast to medaka Admp, whose expression domain is located in the posterior part of the shield at mid-gastrula stage (Fig. 9F"), the amphioxus Admp promoter was activated along the entire shield (Fig. 9F and F'), rather resembling endogenous expression of Admp in amphioxus (Fig. 7G). Finally, we investigated whether the putative regulatory sequences derived from amphioxus Bmp2/4 can be activated in a proper manner in medaka embryo. Interestingly, amphioxus Bmp2/4 and vertebrate Bmp4 genes are structurally organized in a similar manner, possessing the non-coding first exon followed by a relatively long (several kb) intron in front of the first coding exon. Intrigued by this conserved gene structure we prepared a reporter gene construct, designated AmphiBmp2/4-Long-Cherry, which included the first intron in addition to the 5' regulatory region (Fig. 9C). The activity of AmphiBmp2/4-Long-Cherry reporter was stronger in the ventrolateral domain than in the dorsal domain (Fig. 9G and G'), resembling the endogenous expression of medaka Bmp4 (Fig. 9G"). In order to see whether some spatial or temporal regulatory information is indeed located within the first intron we prepared a construct, designated AmphiBmp2/4-Short-Cherry, in which the intronic sequence was deleted, leaving only about 800 base pairs of the 5' regulatory region including the transcriptional start site (Fig. 9D). Interestingly, this truncated reporter construct was activated at mid-gastrula stage in the dorsal domain of medaka embryo, suggesting that the intronic region contains an essential regulatory information required for the expression of amphioxus Bmp2/4 in the ventrolateral domain (Fig. 9H and H'). These results highlight the

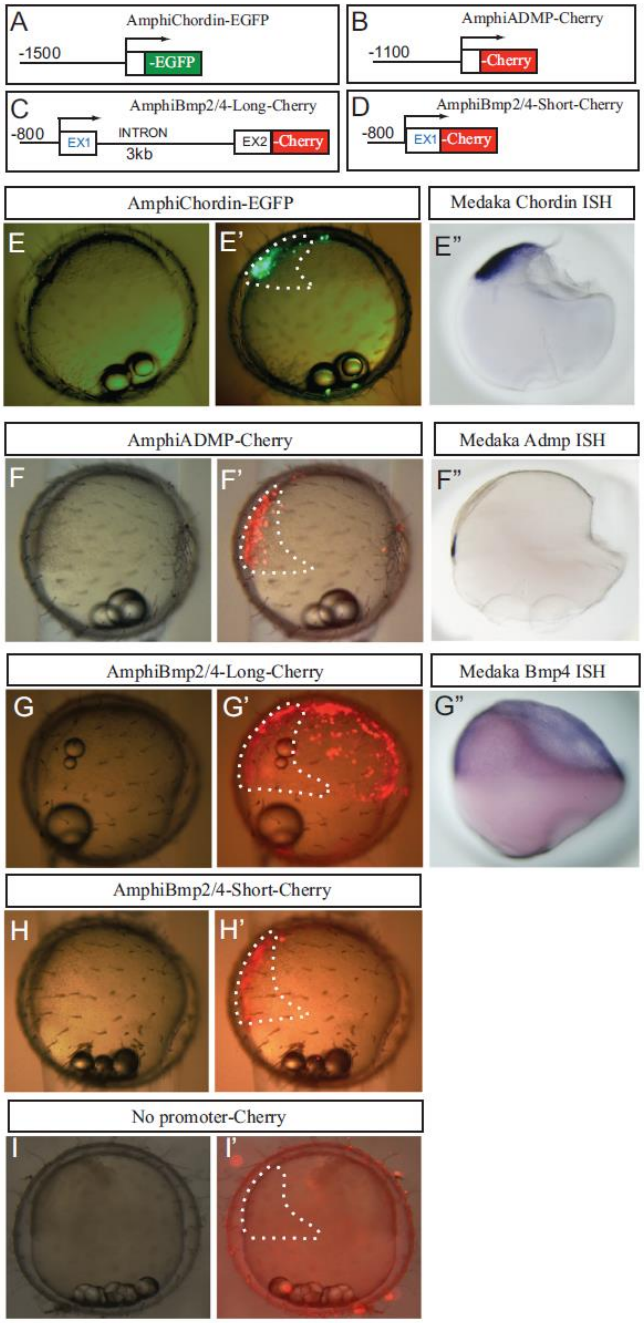


Fig. 9. Analysis of the regulatory sequences of amphioxus Chordin, Admp and Bmp2/4 genes in medaka embryo. Schematic representation of reporter gene constructs, containing amphioxus Chordin (A), Admp (B) and Bmp2/4 (C and D) *cis*-regulatory sequences, that were used for transient transgenesis in the medaka embryo. Transient expression of EGFP in medaka embryos injected with the AmphiChordin-EGFP construct examined under bright field (E) and bright field merged with fluorescence (E′) was seen in the dorsal shield of early medaka gastrula and recapitulated the expression pattern of medaka Chordin (E″). Transient expression of Cherry in medaka embryo injected with AmphiAdmp-Cherry (F and F′), AmphiBmp2/4-Long-Cherry (G and G′) and AmphiBmp2/4-Short-Cherry (H and H′), examined under bright field (F, G, H) and bright field merged with fluorescence (F′, G′, H′). The expression of Cherry driven by the amphioxus Admp *cis*-regulatory sequence was visible throughout the dorsal shield of the embryo at mid-gastrula stage (F and F′). (F″) The expression of endogenous medaka Admp. The AmphiBmp2/4-Long-Cherry reporter gene was active in the ventrolateral domain of the medaka embryo (G and G′) and recapitulated the expression pattern of endogenous medaka Bmp4 gene (G″). In contrast, microinjection of AmphiBmp2/4-Short-Cherry resulted in dorsal expression of Cherry (H and H′). The embryos injected with the promoter-less Cherry reporter gene construct exhibited no fluorescence (I and I′). All embryos are in lateral view. White dashed line indicates position of the shield in medaka embryo.

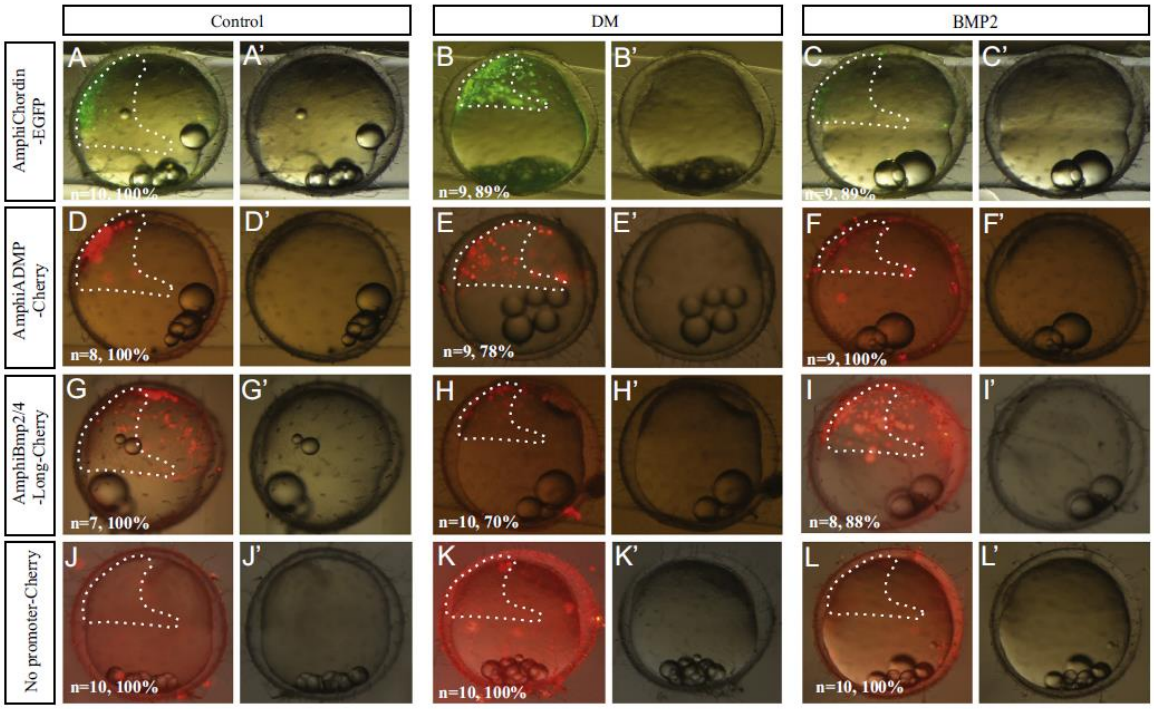


Fig. 10. The embryos injected with AmphiChordin-EGFP (A–C′), AmphiAdmp-Cherry (D–F′), AmphiBmp2/4-Long-Cherry (G–I′) constructs and the promoter-less Cherry containing only the reporter gene (J–L′) were treated with DMSO, 20 μM of DM or 250 ng/ml of recombinant BMP2 at early blastula stage and the activity of the reporter was monitored at mid-gastrula stage. Application of DM led to the increased activity of reporter genes driven by amphioxus Chordin (B and B′) and Admp *cis*-regulatory sequences (E and E') as compared to controls ((A and A′) and (D and D′), respectively). In contrast, the signal of AmphiBmp2/4-Long-Cherry was decreased in the DM treated embryos (H and H′) as compared to controls (G and G′). The treatment with BMP2 led to the downregulation of reporter genes driven by amphioxus Chordin (C and C′) and Admp (F and F′) *cis*-regulatory sequences. Conversely, the activity of AmphiBmp2/4-Long-Cherry was expanded into the dorsal shield of medaka embryo (I and I′). The construct containing the promoter-less Cherry reporter gene was not active in the control and treated embryos (J–L′). White dashed line indicates position of the shield in medaka embryo.

importance of intronic regulatory sequences for correct spatial regulation Bmp2/4 gene and indicate a complexity of Bmp2/4 *cis*-regulatory logic. We further characterized the *cis*-regulatory modules of amphioxus Chordin, Admp and Bmp2/4 genes in the context of vertebrate embryo. The medaka embryos injected with AmphiChordin-EGFP, AmphiAdmp-Cherry and AmphiBmp2/4-Long-Cherry were subsequently treated with BMP2 or DM at early blastula stage and analyzed at mid-gastrula stage. Similarly to endogenous genes in amphioxus (Fig. 3D and D') or medaka (Fig. 6E and F), reporter gene driven by regulatory sequence from amphioxus Chordin was upregulated in the embryos treated with DM (compare Fig. 10A, A' and B, B'). Conversely, the treatment with BMP2 led to the downregulation of the AmphiChordin-EGFP activity (compare Fig. 10A, A' and C, C'). Likewise, the activity of AmphiAdmp-Cherry was increased in DM-treated embryos (compare Fig. 10D D' and E-E') and downregulated in BMP2-treated embryos (compare Fig. 10D D' and F-F'). In the case of AmphiBmp2/4-Long-Cherry construct we observed weakening of the signal in the embryos treated with the inhibitor of Bmp signaling DM as compared to the controls (compare Fig. 10G, G' and H, H'). The treatment with BMP2 resulted in the expansion of the reporter activity driven by amphioxus Bmp2/4 regulatory sequence into the dorsal shield of medaka embryos (compare Fig. 10G, G' and I, I'). We did not observe any signal in the embryos injected with the promoter-less Cherry reporter gene (Fig. 10J–L'). Taken together, our analysis of the gene regulatory sequences of amphioxus Chordin, Admp and Bmp2/4 genes in the context of vertebrate embryo provides another level of evidence for the deep evolutionary conservation of underlying GRN between vertebrates and invertebrate chordates.

Discussion

A positive autoregulatory feedback loop in Bmp signaling as a lockdown of early cell fate specification in axial patterning of bilaterians

Cell fate determination as a function of a spatiotemporal gene expression is thought to be the essential principle of development. Before the cells reach their final commitment they usually pass through many intermediate states, which can be reversible or irreversible. In terms of gene regulatory network the irreversible state is provided by lockdown mechanisms (Davidson, 2010). One of the common lockdown mechanisms is described as a positive feedback loop or dynamic feedback lockdown subcircuit (Davidson, 2009) that provides the irreversibility of the cell fate. To separate multicellular domains within the embryo it is necessary to create a border between two cell fates located side by side. In terms of the GRN logic it is suggested that two irreversible cell states have to repress each other (Peter and Davidson, 2011). It is obvious that such mechanism is present during early development of vertebrate dorsal and ventral domains in the form of Bmp/Chordin network. For the ventral domain specification, Bmp signaling promotes ventral specific and

suppresses organizer specific genes in the vertebrate embryo (Kishimoto et al., 1997; Nguyen et al., 1998; Weaver and Kimelman, 2004). The Bmp positive feedback loop apparently serves as one of the mechanisms providing a high level of Bmp signaling and it plays an important role in the lockdown of cell fate. In zebrafish, both Bmp2b and Bmp7 are required for the maintenance of Bmp2b, Bmp7, and Bmp4 expression in ventral regions of developing embryos (Kishimoto et al., 1997; Schmid et al., 2000). Bmp4 can activate its own expression in *Xenopus* overexpression experiments (Jones et al., 1992). In this study we have shown that inhibition of Bmp signaling by dorsomorphin reduces the level of Bmp4 expression in medaka. Using the same approach we also provide compelling evidence that Bmp signaling is required for the maintenance of Bmp2/4 and Bmp5/8 expression during gastrulation in a basal chordate amphioxus. In addition, inhibition of Bmp signaling leads to the upregulation of dorsal and neural plate specific genes.

It was shown that amphioxus protein Chordin is in biochemical terms functionally equivalent to vertebrate Chordins and that morpholino-based knockdown of Chordin leads to transcriptional upregulation of Bmp2/4 at late gastrula stage (Onai et al., 2010). These data are consistent with our study showing the presence of positive feedback loop in the transcriptional regulation of amphioxus Bmp2/4 and Bmp5/8. The Bmp signaling positive feedback loop may play a key role in the establishment of mutually exclusive Bmp/Chordin expression pattern and, thus, in the formation of the border between dorsal and ventral domains. In the chordate embryo, a border is created between the high and low level of Bmp signaling activity gradient, providing the possibility for ectodermal and endomesodermal domains to adopt distinct cell fates. It is likely that the difference between the two domains and sharpness of the border would be dependent on the strength of additional secondary positive feedback loops. In vertebrates, one additional positive feedback loop exists in the form of Vent genes, which were shown to be direct targets of Bmp signaling in *Xenopus* (Henningfeld et al., 2000; Rastegar et al., 1999) and at the same time activators of Bmp genes (Schuler-Metz et al., 2000). In zebrafish, mutational analysis and loss-of function experiments indicated that Vox and Vent genes are required for maintaining Bmp4 and Bmp2 expression (Imai et al., 2001). Here we show that the upregulation of Bmp4 expression in medaka is stimulated by medaka Vent mRNA microinjection. We have previously shown that similarly to *Xenopus* Xvent-2 gene promoter, promoters of amphioxus Vent1 and Vent2 are Bmp-responsive (Kozmikova et al., 2011). Here we confirmed our previous data by showing decreased expression of Vent1 and Vent2 in amphioxus embryos treated with an inhibitor of Bmp signaling. Our gain-of-function experiments suggest that Vent1 can repress Chordin but is not able to activate Bmp2/4 expression at gastrula and neurula stages of amphioxus. These results imply a modification of GRN architecture in basal chordates as compared to vertebrates (Fig. 11C and D). We hypothesize that the differences in the system of one additional positive loop could result in the establishment of ventrolateral gradient of the Bmp signaling level, which is slightly distinct in vertebrates and basal chordate amphioxus. In particular, since Vent1 is expressed in the neural plate border region in amphioxus (Fig. 3F) and Xvent-2 is expressed in neural plate borders in *Xenopus* (Martynova et al., 2004), this difference in gene regulation, elicited by a change in GRN architecture (such as modification of a subcircuit), might be responsible for the differences in the cell commitment of amphioxus border region *versus* vertebrates. As was shown previously, even small differences in the patterning and level of Bmp signaling have profound effects on cell differentiation in *Drosophila* neural ectoderm border region (Peluso et al., 2011). It is interesting to note that the positive feedback loop of Bmp2/4 is also present in ecdysosoa (Biehs et al., 1996; Mizutani and Bier, 2008), which possess a centralized nervous system.

Reminiscent of the early development of amphioxus, during gastrulation of cnidarian *Nematostella* the two-layered gastrula is developed (Kraus and Technau, 2006; Magie et al., 2007). However, in contrast to amphioxus, the two well-separated domains with neural or non-neural cell fate are not present in *Nematostella* embryo at the end of gastrulation (Marlow et al., 2009). It was shown that the Bmp negative feedback loop is present during early

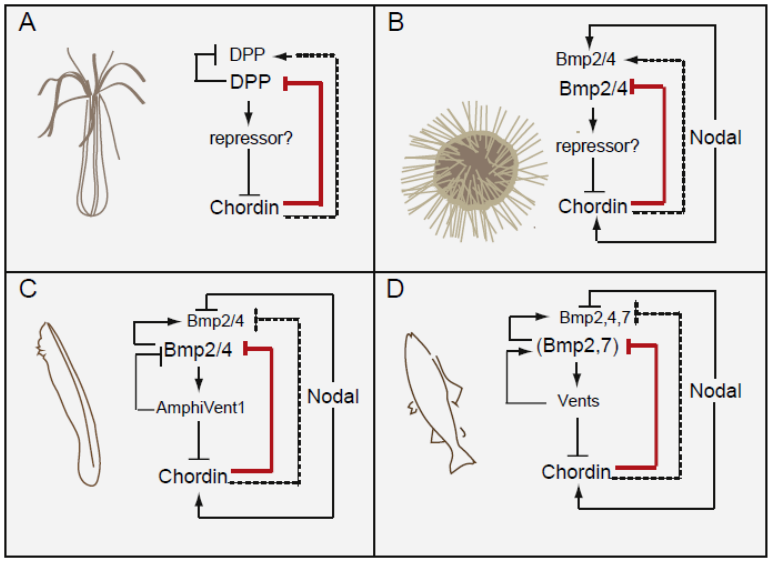


Fig. 11. Simplified schemes depicting Bmp–Chordin GRNs in cnidarian *Nematostella* (A), echinoderm sea urchin (B), cephalochordate amphioxus (C) and vertebrate zebrafish (D) during gastrulation. Red lines from Chordin mark antagonism of Bmp signaling at the protein level. Black lines denote transcriptional regulatory links. Dashed lines indicate an indirect effect of Chordin on the regulation of Dpp/Bmp transcription that results in activation in the case of *Nematostella* (A) and sea urchin (B) and repression in the case of chordate embryos (C and D).

Nematostella development. As a consequence, *Nematostella* Chordin, which is an inhibitor of Bmp signaling at the protein level (Rentzsch et al., 2006), works as a transcriptional activator of Dpp in a double negative loop (Saina et al., 2009) (Fig. 11A). Furthermore, it is obvious that any anti-Bmp secreted protein modulator will have the same effect on the transcriptional readout in *Nematostella* GRN. In terms of GRN logic, activation of Dpp and Bmp5/8 by Chordin might be responsible for their partial co-expression. Such conditions do not allow cells to separate their fate early into two distinct domains, which results in the absence of a clear division of the ectoderm into neural and non-neural during gastrulation of *Nematostella*. We therefore speculate that during evolution the emergence of the Bmp signaling positive feedback loop led to the establishment of a molecular platform (GRN architecture) for the development of centralized nervous system (Fig. 11A, C and D).

The nervous system of another basal deuterostome, sea urchin, has the nerves organized into a highly restricted pattern within the ectoderm forming the so-called ciliary band ring. It has also been argued that the nervous system of sea urchin is diffuse during development (Angerer et al., 2011). Similarly to vertebrates and amphioxus, one of the main conditions required for establishment of neural cell fate in the ciliary band is a low level of Bmp signaling (Angerer et al., 2000; Bradham et al., 2009). In contrast to chordates and insects, sea urchin Bmp2/4 does not activate itself during early development (Fig. 11B) (Saudemont et al., 2010). As in the case of *Nematostella*, sea urchin Chordin and Bmp2/4 are co-expressed in the ventral ectoderm of the embryo that is consistent with a proposed GRN logic (Bradham et al., 2009; Lapraz et al., 2009; Saudemont et al., 2010). Although additional specialized lineage-specific mechanisms are utilized in the sea urchin embryo to translocate Bmp signaling activity to the dorsal domain of the embryo (Lapraz et al., 2009), the most ventral domain, which serves as a source of Bmp2/4, is not fated into neural fate (Saudemont et al., 2010). The co-expression of Chordin and Bmp2/4 in the ventral ectoderm of sea urchin is also promoted by Nodal, which is not required for the expression of amphioxus Bmp2/4 (this study) and was not found in *Nematostella*. It seems that the positioning of the nervous system dorsally in chordates depends greatly on the factors that suppress BMP signaling, suggesting a key role of the positive feedback loop in the process of centralization. To better understand the evolution of nervous system centralization it would be interesting to find out what initial signals are required for Bmp2/4, Bmp5/8 and Chordin expression in the early cnidarian embryo.

Conservation of cis-regulatory sequences in genes patterning early chordate embryo

Cis-regulatory sequences represent the summary of regulatory information required for the spatio-temporal pattern of the corresponding genes. That is why cross-species transgenesis of the regulatory elements provides the information about the conservation of several regulatory inputs and together with comparison of expression patterns may serve as a useful approach to identify the ancestral characters. Recently, some data were published revealing deep conservation of *cis*-regulatory sequences (Ochi et al., 2012; Royo et al., 2011). It is interesting to note that the regulatory regions identified in amphioxus so far have been localized close to the 5'end of the gene (Beaster-Jones et al., 2007; Kozmikova et al., 2011; Yu et al., 2004). Here we found that only short 5' regulatory regions of amphioxus Chordin, Admp and Bmp2/4 are sufficient for correct spatio-temporal activation in the medaka early embryo. It is likely that due to a smaller genome size of amphioxus as compared to vertebrates its *cis*-regulatory regions are generally located in close proximity to the corresponding genes. We have shown that the *cis*-regulatory elements of amphioxus organizer specific genes Chordin and Admp drive the expression of the reporters into the organizer region of medaka embryo. We observed that the spatial activity of the *cis*-regulatory sequence of amphioxus Bmp2/4 is largely similar to that of endogenous Bmp4 and the spatial activity of the reporter gene driven by amphioxus Chordin is identical with the expression pattern of its endogenous counterpart in medaka organizer. These data suggest remarkable conservation of regulatory input of both these genes. The activity of the amphioxus Admp promoter in the anterior and posterior dorsal midline is more reminiscent of endogenous expression of the gene in amphioxus embryo. Likewise, the promoter of amphioxus FoxD was shown to be active in the chicken embryo in similar regions, where the gene is expressed in amphioxus but not in the expression domains of endogenous chicken FoxD (Yu et al., 2002). It is important to note that the expression pattern of medaka Admp, which is detected in the posterior dorsal domain of the shield, is different from the amphioxus, *Xenopus* and zebrafish counterparts, which are all expressed in anterior and posterior dorsal midline (Lele et al., 2001; Moos Jr. et al., 1995; Yu et al., 2007). Our study demonstrated that the regulation of the *cis*-regulatory sequences of amphioxus Bmp2/4, Chordin and Admp genes with Bmp signaling in the context of medaka embryo is reminiscent of the regulation of endogenous genes in medaka and amphioxus. Thus, we provide additional evidence that the mechanisms underlying axial patterning of amphioxus and vertebrate gastrula are highly evolutionarily conserved. Further detailed investigations of GRN together with cell fate tracing in the cephalochordate embryo and other deuterostomes may uncover the ancestor state and at the same time reveal some features that are responsible for the differences in the establishment of cell fates in diverse animal lineages.

Acknowledgments

We are grateful to Hector Escriva and Stephanie Bertrand, Laboratoire Arago, Banyuls sur Mer, France, for providing excellent logistic and laboratory help during the summer breeding season of *B. lanceolatum*. We thank Susan Bell of the University of South Florida, Tampa, Fl, USA, for providing laboratory space during the summer breeding season of *B. floridae*. We are grateful to Sarka Takacova for manuscript proofreading. This study was supported by the Grant Agency of the Czech Republic (P305/10/J064) and by OP EC (CZ1.07/2.3.00/30.0027 "Founding the Centre of Transgenic Technologies" Operational Program Ministry of Education, Youth and Sports, Czech Republic+European Social Fund) to Z.K., by the ASSEMBLE Program at Laboratoire Arago, Banyuls sur Mer (ASSEMBLE grant agreement no. 227799) to I.K., by IMG Institutional Support RVO68378050, and by MIUR (PRIN 20088JEHW3-65 001) to S.C.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.ydbio.2013.07.021.

References

Angerer, L.M., et al., 2000. A BMP pathway regulates cell fate allocation along the sea urchin animal–vegetal embryonic axis. Development 127, 1105–1114.
Angerer, L.M., et al., 2011. The evolution of nervous system patterning: insights from sea urchin development. Development 138, 3613–3623.
Arendt, D., Nubler-Jung, K., 1994. Inversion of dorsoventral axis? Nature 371, 26.
Ashe, H.L., et al., 2000. Dpp signaling thresholds in the dorsal ectoderm of the Drosophila embryo. Development 127, 3305–3312.

Bayascas, J.R., et al., 2002. Isolation of AmphiCASP-3/7, an ancestral caspase from amphioxus (*Branchiostoma floridae*). Evolutionary considerations for vertebrate caspases. Cell Death Differ. 9, 1078–1089.

Beaster-Jones, L., et al., 2007. *Cis*-regulation of the amphioxus engrailed gene: insights into evolution of a muscle-specific enhancer. Mech. Dev. 124, 532–542.

Bertrand, S., Escriva, H., 2011. Evolutionary crossroads in developmental biology: amphioxus. Development 138, 4819–4830.

Biehs, B., et al., 1996. The Drosophila short gastrulation gene prevents Dpp from autoactivating and suppressing neurogenesis in the neuroectoderm. Genes Dev. 10, 2922–2934.

Bradham, C.A., et al., 2009. Chordin is required for neural but not axial development in sea urchin embryos. Dev. Biol. 328, 221–233.

Candiani, S., et al., 2002. Cloning and developmental expression of AmphiBrn1/2/4, a POU III gene in amphioxus. Mech. Dev. 116, 231–234.

Dale, L., et al., 1992. Bone morphogenetic protein 4: a ventralizing factor in early Xenopus development. Development 115, 573–585.

Davidson, E.H., 2009. Network design principles from the sea urchin embryo. Curr. Opinion Genet. Dev. 19, 535–540.

Davidson, E.H., 2010. Emerging properties of animal gene regulatory networks. Nature 468, 911–920.

Denes, A.S., et al., 2007. Molecular architecture of annelid nerve cord supports common origin of nervous system centralization in bilateria. Cell 129, 277–288.

Dosch, R., Niehrs, C., 2000. Requirement for anti-dorsalizing morphogenetic protein in organizer patterning. Mech. Dev. 90, 195–203.

Flores, M.V., et al., 2008. Osteogenic transcription factor Runx2 is a maternal determinant of dorsoventral patterning in zebrafish. Nat. Cell Biol. 10, 346–352.

Fuentes, M., et al., 2007. Insights into spawning behavior and development of the European amphioxus (*Branchiostoma lanceolatum*). J. Exp. Zool. B Mol. Dev. Evol. 308, 484–493.

Gao, H., et al., 2007. Xom interacts with and stimulates transcriptional activity of LEF1/TCFs: implications for ventral cell fate determination during vertebrate embryogenesis. Cell Res. 17, 345–356.

Henningfeld, K.A., et al., 2000. Smad1 and Smad4 are components of the bone morphogenetic protein-4 (BMP-4)-induced transcription complex of the Xvent-2B promoter. J. Biol. Chem. 275, 21827–21835.

Holland, L.Z., et al., 2008. The amphioxus genome illuminates vertebrate origins and cephalochordate biology. Genome Res. 18, 1100–1111.

Holland, L.Z., et al., 2000. Evolutionary conservation of the presumptive neural plate markers AmphiSox1/2/3 and AmphiNeurogenin in the invertebrate chordate amphioxus. Dev. Biol. 226, 18–33.

Hoppler, S., Moon, R.T., 1998. BMP-2/-4 and Wnt-8 cooperatively pattern the Xenopus mesoderm. Mech. Dev. 71, 119–129.

Imai, Y., et al., 2001. The homeobox genes vox and vent are redundant repressors of dorsal fates in zebrafish. Development 128, 2407–2420.

Iwamatsu, T., 2004. Stages of normal development in the medaka *Oryzias latipes*. Mech. Dev. 121, 605–618.

Jones, C.M., et al., 1992. DVR-4 (bone morphogenetic protein-4) as a posterior-ventralizing factor in Xenopus mesoderm induction. Development 115, 639–647.

Kim, J., et al., 1998. Transcriptional regulation of BMP-4 in the Xenopus embryo: analysis of genomic BMP-4 and its promoter. Biochem. Biophys. Res. Commun. 250, 516–530.

Kishimoto, Y., et al., 1997. The molecular nature of zebrafish swirl: BMP2 function is essential during early dorsoventral patterning. Development 124, 4457–4466.

Kozmik, Z., et al., 2001. Characterization of Amphioxus AmphiVent, an evolutionarily conserved marker for chordate ventral mesoderm. Genesis 29, 172–179.

Kozmik, Z., et al., 2007. Pax–Six–Eya–Dach network during amphioxus development: conservation *in vitro* but context specificity *in vivo*. Dev. Biol. 306, 143–159.

Kozmikova, I., et al., 2011. Conservation and diversification of an ancestral chordate gene regulatory network for dorsoventral patterning. PloS One 6, e14650.

Kraus, Y., Technau, U., 2006. Gastrulation in the sea anemone *Nematostella vectensis* occurs by invagination and immigration: an ultrastructural study. Dev. Genes Evol. 216, 119–132.

Kwon, H.J., et al., 2010. Identification of early requirements for preplacodal ectoderm and sensory organ development. PLoS Genet 6, e1001133.

Lapraz, F., et al., 2009. Patterning of the dorsal–ventral axis in echinoderms: insights into the evolution of the BMP-chordin signaling network. PLoS Biol. 7, e1000248.

Lele, Z., et al., 2001. Zebrafish admp is required to restrict the size of the organizer and to promote posterior and ventral development. Dev. Dyn. 222, 681–687.

Lowe, C.J., et al., 2006. Dorsoventral patterning in hemichordates: insights into early chordate evolution. PLoS Biol. 4, e291.

Lu, T.M., et al., 2012. BMP and Delta/Notch signaling control the development of amphioxus epidermal sensory neurons: insights into the evolution of the peripheral sensory system. Development 139, 2020–2030.

Magie, C.R., et al., 2007. Gastrulation in the cnidarian *Nematostella vectensis* occurs via invagination not ingression. Dev. Biol. 305, 483–497.

Marchal, L., et al., 2009. BMP inhibition initiates neural induction via FGF signaling and Zic genes. Proc. Natl. Acad. Sci. USA 106, 17437–17442.

Marlow, H.Q., Srivastava, M., Matus, D.Q., Rokhsar, D., Martindale, M.Q., 2009. Anatomy and development of the nervous system of Nematostella vectensis, an anthozoan cnidarian. Dev. Neurobiol. 69, 235–254.

Martynova, N., et al., 2004. Patterning the forebrain: FoxA4a/Pintallavis and Xvent2 determine the posterior limit of Xanf1 expression in the neural plate. Development 131, 2329–2338.

Matus, D.Q., et al., 2006. Dorso/ventral genes are asymmetrically expressed and involved in germ-layer demarcation during cnidarian gastrulation. Curr. Biol. 16, 499–505.

Metz, A., et al., 1998. Structural and functional analysis of the BMP-4 promoter in early embryos of *Xenopus laevis*. Mech. Dev. 74, 29–39.

Meulemans, D., Bronner-Fraser, M., 2007. The amphioxus SoxB family: implications for the evolution of vertebrate placodes. Int. J. Biol. Sci. 3, 356–364.

Mizutani, C.M., Bier, E., 2008. EvoD/Vo: the origins of BMP signalling in the neuroectoderm. Nat. Rev. Genet. 9, 663–677.

Mizutani, C.M., et al., 2006. Threshold-dependent BMP-mediated repression: a model for a conserved mechanism that patterns the neuroectoderm. PLoS Biol. 4, e313.

Moos Jr., M., et al., 1995. Anti-dorsalizing morphogenetic protein is a novel TGF-beta homolog expressed in the Spemann organizer. Development 121, 4293–4301.

Nguyen, V.H., et al., 1998. Ventral and lateral regions of the zebrafish gastrula, including the neural crest progenitors, are established by a bmp2b/swirl pathway of genes. Dev. Biol. 199, 93–110.

Ochi, H., et al., 2012. Evolution of a tissue-specific silencer underlies divergence in the expression of pax2 and pax8 paralogues. Nat. Commun. 3, 848.

Onai, T., et al., 2010. Opposing Nodal/Vg1 and BMP signals mediate axial patterning in embryos of the basal chordate amphioxus. Dev. Biol. 344, 377–389.

Onichtchouk, D., et al., 1996. The Xvent-2 homeobox gene is part of the BMP-4 signalling pathway controlling [correction of controling] dorsoventral patterning of Xenopus mesoderm. Development 122, 3045–3053.

Onichtchouk, D., et al., 1998. Requirement for Xvent-1 and Xvent-2 gene function in dorsoventral patterning of Xenopus mesoderm. Development 125, 1447–1456.

Peluso, C.E., et al., 2011. Shaping BMP morphogen gradients through enzyme-substrate interactions. Dev. Cell 21, 375–383.

Peter, I.S., Davidson, E.H., 2011. Evolution of gene regulatory networks controlling body plan development. Cell 144, 970–985.

Rankin, S.A., et al., 2011. A gene regulatory network controlling hhex transcription in the anterior endoderm of the organizer. Dev. Biol. 351, 297–310.

Rastegar, S., et al., 1999. Transcriptional regulation of Xvent homeobox genes. Mech. Dev. 81, 139–149.

Rembold, M., Lahiri, K., Foulkes, N.S., Wittbrodt, J., 2006. Transgenesis in fish: efficient selection of transgenic fish by co-injection with a fluorescent reporter construct. Nat. Protoc. 1, 1133–1139.

Rentzsch, F., et al., 2006. Asymmetric expression of the BMP antagonists chordin and gremlin in the sea anemone *Nematostella vectensis*: implications for the evolution of axial patterning. Dev. Biol. 296, 375–387.

Reversade, B., De Robertis, E.M., 2005. Regulation of ADMP and BMP2/4/7 at opposite embryonic poles generates a self-regulating morphogenetic field. Cell 123, 1147–1160.

Royo, J.L., et al., 2011. Transphyletic conservation of developmental regulatory state in animal evolution. Proc. Natl. Acad. Sci. USA 108, 14186–14191.

Saina, M., et al., 2009. BMPs and chordin regulate patterning of the directive axis in a sea anemone. Proc. Natl. Acad. Sci. USA 106, 18592–18597.

Saudemont, A., et al., 2010. Ancestral regulatory circuits governing ectoderm patterning downstream of Nodal and BMP2/4 revealed by gene regulatory network analysis in an echinoderm. PLoS Genet. 6, e1001259.

Schmid, B., et al., 2000. Equivalent genetic roles for bmp7/snailhouse and bmp2b/swirl in dorsoventral pattern formation. Development 127, 957–967.

Schuler-Metz, A., et al., 2000. The homeodomain transcription factor Xvent-2 mediates autocatalytic regulation of BMP-4 expression in Xenopus embryos. J. Biol. Chem. 275, 34365–34374.

von Ohlen, T., Doe, C.Q., 2000. Convergence of dorsal, dpp, and egfr signaling pathways subdivides the Drosophila neuroectoderm into three dorsal–ventral columns. Dev. Biol. 224, 362–372.

Weaver, C., Kimelman, D., 2004. Move it or lose it: axis specification in Xenopus. Development 131, 3491–3499.

Willot, V., et al., 2002. Cooperative action of ADMP- and BMP-mediated pathways in regulating cell fates in the zebrafish gastrula. Dev. Biol. 241, 59–78.

Yu, J.K., Holland, L.Z., 2009a. Amphioxus (*Branchiostoma floridae*) spawning and embryo collection. Cold Spring Harbor Protocol 9, http://dx.doi.org/10.1101/pdb.prot5285.

Yu, J.K., Holland, L.Z., 2009b. Amphioxus whole-mount *in situ* hybridization. Cold Spring Harbor Protocol 9, http://dx.doi.org/10.1101/pdb.prot5286.

Yu, J.K., et al., 2002. An amphioxus winged helix/forkhead gene, AmphiFoxD: insights into vertebrate neural crest evolution. Dev. Dyn. 225, 289–297.

Yu, J.K., et al., 2004. Tissue-specific expression of FoxD reporter constructs in amphioxus embryos. Dev Biol 274, 452–461.

Yu, J.K., et al., 2008. Insights from the amphioxus genome on the origin of vertebrate neural crest. Genome Res. 18, 1127–1132.

Yu, J.K., et al., 2007. Axial patterning in cephalochordates and the evolution of the organizer. Nature 445, 613–617.

Yu, P.B., et al., 2008. Dorsomorphin inhibits BMP signals required for embryogenesis and iron metabolism. Nat. Chem. Biol. 4, 33–41.

VII.2.2. **Fabian P, Pantzartzi CN, Kozmikova I, Kozmik Z.** *vox* homeobox gene: a novel regulator of midbrain-hindbrain boundary development in medaka fish. **Dev Genes Evol. 2016**

vox homeobox gene: a novel regulator of midbrain-hindbrain boundary development in medaka fish?

Peter Fabian[1], Chrysoula N. Pantzartzi[1], Iryna Kozmikova[1], Zbynek Kozmik[1]*

Department of Transcriptional Regulation, Institute of Molecular Genetics, Videnska 1083, Prague 14220, Czech Republic

*Corresponding author:
Dr. Zbynek Kozmik
E-mail: kozmik@img.cas.cz

[1] Institute of Molecular Genetics of the AS CR, v. v. i., Videnska 1083, 14220 Prague 4, Czech Republic.

ABSTRACT

The midbrain-hindbrain boundary (MHB) is one of the key organizing centers of the vertebrate central nervous system (CNS). Its patterning is governed by a well-described gene regulatory network (GRN) involving several transcription factors, namely *pax*, *gbx*, *en* and *otx,* together with signaling molecules of the Wnt and Fgf families. Here we describe the onset of these markers in *Oryzias latipes* (medaka) early brain development in comparison to previously known zebrafish expression patterns. Moreover, we show for the first time that *vox*, a member of the vent gene family, is expressed in the developing neural tube similarly to CNS markers. Overexpression of *vox* leads to profound changes in the gene expression patterns of individual components of MHB-specific GRN, most notably of *fgf8,* a crucial organizer molecule of MHB. Our data suggest that genes from the vent family, in addition to their crucial role in body axis formation, may play a role in regionalizing of vertebrate CNS.

KEYWORDS Midbrain-hindbrain boundary, *vox*, medaka, Heat Shock Element, *fgf8*, Gene Regulatory Network

INTRODUCTION

Induction of the neural plate is considered the initial step in the development of the central nervous system (CNS). Patterning of the primordial brain begins shortly after gastrulation, when the neural plate begins to form. During the neurulation process of teleost fish the neural plate firstly transforms into a transient structure called neural keel, which later gives rise to the neural rod (approximately at 5-somite stage). Finally, the neural tube forms at 10-somite stage. The neural keel is regionalized both across the dorso-ventral (D/V) axis and across the antero-posterior (A/P) axis. Later on, the neural rod is divided into three primary brain regions: the forebrain, the midbrain and the hindbrain, as a result of intensive signaling cross-talk. D/V patterning of the neural rod (neural tube) is controlled via ventralizing signals – SHH from notochord cells - and via dorsalizing signals – mainly BMP4 from epidermal ectoderm (Echelard et al. 1993; Lee and Jessell 1999). A/P patterning, on the other hand, is organized by local signaling centers (Echevarria et al. 2003; Jessell and Sanes 2000). One of them is a group of cells at the midbrain-hindbrain boundary (MHB) also called the isthmic organizer (Houart et al. 1998; Wilson and Houart 2004; Wurst and Bally-Cuif 2001). MHB is both required and sufficient for the cell fate induction of the surrounding tissues, which are later developed into mesencephalic and metencephalic regions (Liu and Joyner 2001; Raible and Brand 2004; Wurst and Bally-Cuif 2001). This junction is established, induced and maintained by the expression of several transcription factors, most notably *otx*, *gbx*, *pax* and *en*, and signaling molecules from Wnt and Fgf families. These factors have been found in all studied vertebrate species and are often regarded as the core MHB cascade (reviewed in Dworkin and Jane (2013) and Wurst and Bally-Cuif (2001)). In zebrafish, establishment of MHB is determined at the borders of the expression domains of homeobox transcription factors *otx2* (presumptive midbrain) and *gbx1* (presumptive hindbrain). Positioning of the MHB primordium (*otx2*/*gbx1* borders) in the zebrafish neural plate is mediated by WNT8 proteins from the precursors of lateral mesendoderm (Rhinn et al. 2009; Rhinn et al. 2005). Once the expression of *otx2*/*gbx1* is established, the induction step of MHB with three parallel signaling pathways involving *pax2.1*, *wnt1* and *fgf8* is triggered (Canning et al. 2007; Lun and Brand 1998; Rhinn and Brand 2001). Slightly later, *en2, en3* and *en1*, and *pax2.2, pax5* and *pax8* are expressed across the boundary (Lun and Brand 1998). After the initiation phase, *fgf8* expression is positively regulated by *gbx2,* while *otx2* and *gbx2*/*fgf8* regulate each other negatively, and these interactions lead to the establishment and maintenance of sharp expression borders (Rhinn et al. 2003). Concomitantly, the expression regions of *fgf8*, *wnt1*, *en1/2/3*, *pax2.1/2.2/5/8* become interdependent and establish a positive regulatory loop that is required for maintaining the midbrain-hindbrain identity (Dworkin and Jane 2013; Wurst and Bally-Cuif 2001). The maintenance of MHB relies on continuous expression of core MHB factors: *fgf8*, *wnt1*, *en1/2*, *pax2.1* (Dworkin and Jane 2013; Joyner et al. 2000; Rhinn and Brand 2001; Wurst and Bally-Cuif 2001). Perturbation of any of these genes has no crucial effect on the induction, however, it will lead to severe defects in the maintenance of the isthmic organizer (Dworkin and Jane 2013; Joyner et al. 2000; Rhinn and Brand 2001; Wurst and Bally-Cuif 2001).

Vent family genes, including *vox*, belong to a group of homeobox transcription factors that play a significant role during early development. Overexpression of these genes leads to strong ventralization of zebrafish embryos (Gilardelli et al. 2004). v*ox* and other members of the vent family act as transcriptional repressors (Gilardelli et al. 2004; Kozmikova et al. 2013; Zhao et al. 2013). The role of the *vox* gene (or vent family genes) has been relatively well studied during early development, nevertheless, it is largely unknown during later developmental stages. Our aim was to investigate the role of the *vox* homeobox gene in the brain development. We specifically address the relationship between *vox* and the MHB program. We used *Oryzias latipes* (medaka), an important fish model in studies of vertebrate developmental genetics and comparative genomics. We performed synteny and sequence conservation analysis to validate the identity of the medaka *vox* ortholog. We described its expression pattern during embryogenesis and placed it into the framework of known brain markers. Using a heat-shock inducible *vox* line of medaka we show that *vox* interferes with the gene regulatory network during the maintenance stage of developing MHB. Based on our data, we propose a novel function for the *vox* gene as a negative regulator of *fgf8*, a pivotal organizer molecule of MHB.

MATERIALS AND METHODS

In silico analysis of fish *vox* loci

Genomic loci containing fish *vox* genes were retrieved through ENSEMBL and NCBI websites. Comparison of *vox* syntenic regions was done using the Ensembl and NCBI genome browsers, along with information from the Genomicus website v81.01 (Louis et al. 2015). Protein sequences were aligned using Clustal Omega (Sievers et al. 2011) and visualized using GeneDoc (Nicholas et al. 1997). Accession numbers for *vox* orthologs are medaka (XP_004076914.1, ENSORLP00000000781), fugu (XP_003963708, ENSTRUP00000027430), stickleback (ENSGACP00000003759) and zebrafish (NP_571773, ENSDARP00000131684).

DNA constructs

The gfp:HSE:luciferase vector (Bajoghli et al. 2004) was modified by replacing the *gfp* and luciferase genes with the mCherry cDNA and *vox* open reading frame, respectively, leading to the final mCherry:HSE:*vox* construct.
Probes for *otx2*, *gbx2*, *wnt1*, *fgf8*, *pax2*.2 and *en2* were kindly provided by Dr. Thomas Czerny. As previously described, medaka *pax2.1* is not fully annotated in public databases; therefore, available EST and genomic scaffold sequences were used in order to design primers corresponding to the 5′ UTR of *pax2.1* gene (Fabian et al. 2015). Primers for *gbx1* were designed based on genomic sequence ENSORLG00000017521 and for *vox* on ENSORLG00000000633. Probes were amplified from total cDNA using primers 5′-CCACCCCTATATCTGATTGAACC-3′ and 5′-GTTTCGCCTTGGTGATGCA-3′ for *pax2.1*; 5′-GTGAGGAACCCCAAAATCGT-3′ and 5′-CAGCCACAACTCATAACTTCTCC-3′ for *gbx1*. 5′-ATGGTCAAATACTTTTCAGTAGACT-3′, 5′-TCAGAAGAAATGTTGGTAGTGGAT-3′ for *vox*.
To test morpholino knockdown efficiency, the vox AUG region (the region of vox containing the translation initiation codon, flanked by 12 and 10 bp upstream and downstream, respectively) was fused in frame with the 5′ open reading frame of eGFP and cloned into pCS2+ (Fig. S2a) using primers 5′-CGCGGATCCGCTGGACTGCAAAGATGGTCAAATACTTTGTGAGCAAGGGCGAGGAGCTGTTCA-3′, 5′-CCGCTCGAGTTACTTGTACAGCTCGTCCATGCCGAGA-3′ leading to the final vox_aug_region-eGFP construct.

Microinjection of mRNA and morpholino oligonucleotides into medaka eggs

The antisense translation blocking morpholino (MO) against *vox* gene was designed and manufactured by Gene Tools (Philomath, OR, USA). The *vox* MO sequence was 5′-AGTATTTGACCATCTTTGCAGTCCA-3′. Morpholino against zebrafish *tcf7l1b* 5′-CGCCTCCGTTAAGCTGCGGCATGTT-3′ was used as a MO injection control (control MO). Morpholinos were dissolved to 1 mM injection stocks in RNase-free water. RNA for injection was prepared using the mMessage mMachine kit (Ambion). The vox aug region-eGFP construct was subcloned into the pCS2+ expression vector to produce mRNA using SP6 RNA polymerase. For microinjection of medaka, a solution of 1xERM, 0.5% of Phenol Red, 0.3mM morpholino and 200ng/µl of synthetized mRNA was applied.
The injecting setup was as follows: pressure injector Femtojet (Eppendorf); micromanipulator TransferMan NK (Eppendorf); borosilicate glass capillaries (GC100F10, Harward Apparatus); stereomicroscopes (Olympus SZX7, SZX9). Pressure was adjusted so that a bolus of RNA of about 2 nl volume was injected into one-cell stage medaka. Calibration was done with the Stage Micrometer. Injected embryos were fixed at the appropriate stage.

Animal husbandry

Oryzias latipes (medaka) embryos of the Cab inbred strain were used for all experiments. Embryos were collected daily immediately after spawning. Embryonic stages were determined according to

Iwamatsu (2004). Housing of animals and *in vivo* experiments were performed after approval by the Animal Care Committee of the Institute of Molecular Genetics (study ID#36/2007) and in compliance with national and institutional guidelines (ID#12135/2010-17210).

I-SceI meganuclease transgenesis

To generate transgenic lines, the mCherry:HSE:*vox* construct was injected (see paragraph Microinjection of mRNA and morpholino oligonucleotides into medaka eggs) at a concentration of 10 ng/µl together with I-SceI meganuclease (0.5 units/µl) into embryos at the one-cell stage, as described by Thermes et al. (2002) and Bajoghli et al. (2004). The mature F0 fish were crossed with wild-type fish and their F1 progeny was assayed for transgene expression after heat shock. More specifically, F1 offspring at stage 16 (75% epiboly) was treated at 39°C for 2h and fixed 24h later (stage 20-21).

Whole-mount RNA *in situ* hybridization

Embryos were fixed overnight in 4% PFA/PTW (PBS/Tween) at RT and subsequently dechorionated manually. Whole-mount RNA *in situ* hybridizations were performed at 65°C using DIG or fluorescein-labeled RNA probes. The color reaction was carried out with BM purple (Roche) and ImmPACT Vector Red (Vector), followed by re-fixation in 4% PFA/PTW. Embryos were bleached in 100% MetOH, rehydrated in PTW and mounted in 87% glycerol. Whole-mount specimens were cryoprotected in 30% sucrose/PBS, and embedded and frozen in OCT (Tissue Tek; Sakura Finetek). Serial frozen sections were prepared at 10 µm thickness.

RESULTS AND DISCUSSION

In silico analysis of fish *vox* loci

Single *vox* loci were obtained for the four fish species analyzed in this study (Fig. 1a). In particular, *vox* genes are located on chromosomes 15 (NC_019873.1), 4 (NC_018893.1) and 13 (NC_007124.6) of medaka, fugu and zebrafish, respectively, and groupVI (ENSGACG00000002870) of the stickleback genome assembly. Medaka *vox* locus exhibits higher conservation of the gene content and order to the syntenic region of fugu and stickleback than to that of zebrafish. For example, a local rearrangement of the *dntt* gene is observed in zebrafish in relation to the other species.

vox genes of all species compared in this study consist of three exons with comparable size and conserved borders (data not shown). The medaka *vox* ortholog consists of 244 aminoacids. Protein alignment reveals a high degree of conservation for all four species, mainly in the homeobox domain (Fig. 1b, black line). It must be noted that medaka *vox* is mistakenly annotated as "homeobox protein vent1-like" in GenBank (XP_004076914.1), nevertheless, both synteny and sequence alignment support the fact that it is indeed the *vox* ortholog.

Besides the homeobox domain (Fig. 1b, black line), highly conserved blocks are observed in the N-terminus (Fig. 1b, blue and red lines). Previous studies have shown that the *vox*/*vent* family members bear the Eh1 motif (Engrailed Homolog 1) in the N-terminus, a domain that is involved in protein–protein interactions and acts as a repressor in endoderm development (Gilardelli et al. 2004; Zhao et al. 2013). This motif also appears to be present in the medaka ortholog (Fig. 1b, blue line).

Expression of medaka *vox*

To determine the spatial distribution of *vox* transcripts, whole mount RNA *in situ* hybridization was performed with medaka embryos at stages 11-25 (Fig. 2). *vox* transcripts are detectable after zygotic genome activation during the late blastula stage (Fig. 2a), throughout the whole embryo in a mottled pattern. Similarly, in zebrafish the onset of *vox* expression is at the sphere stage (Melby et al. 2000). Later, during gastrulation, *vox* expression is strong at the blastoderm margin and flanks the dorsal side of the embryo, which corresponds to the organizer and neuroectoderm (Fig. 2b, c). *vox* is strongly expressed at the embryonic ventral part - outside of the shield domain - in Fig. 2c, same as *bmp4* expression pattern (Kozmikova et al. 2013). We have previously shown that medaka *vox* is a ventralizing factor (Kozmikova et al. 2013), similarly to vent family genes in zebrafish that suppresses the activity of dorsal genes such as chordin (Melby et al. 2000). By stage 15 (50% epiboly), *vox*

surrounds the presumptive neural plate (Fig. 2b, Fig. S1). Although *vox* is known as a negative regulator of neural tissues, it exhibits a striped expression pattern in the neural plate at the level of prospective midbrain (Fig. 2c, Fig. S1). A similar expression pattern is observed for the markers of the central nervous system (CNS) (Fig. 3). At stage 17 (early neurula), *vox* transcripts become restricted to dorsal side of neural keel at the level of prospective midbrain (Fig. 2d). Double whole-mount RNA *in situ* hybridization revealed partially overlapping expression (displayed with red arrowhead in Fig. 2e, e´) of *vox* with *pax2.2*, a marker of the MHB region (Fig. 2e, e´) (Lun and Brand 1998) in dorsal part of the *pax2.2* expression domain. From the stage 17 (early neurula) through somitogenesis, *vox* is strongly expressed posteriorly in the tail bud, similarly to zebrafish (Gilardelli et al. 2004; Melby et al. 2000). Strong expression of *vox* is maintained in the tail bud (Fig. 2f, i), eye primordia (Fig. 2f, g, i) and in the neural tube within the midbrain region (Fig. 2f-h) (Gilardelli et al. 2004). Collectively, the expression of *vox* in medaka and other vertebrates indicates a role for this gene in CNS development.

Midbrain-hindbrain boundary markers in medaka

We sought to investigate the temporal and spatial distribution of early CNS markers in medaka embryos and compare it to that of zebrafish (Fig. 3a-j). Medaka *otx2* as well as zebrafish *otx2* (Li et al. 1994) are expressed in the most anterior domain, in the prospective neural plate starting from mid-gastrula until the beginning of neurulation (Fig. 3a, Fig. S1). The expression domains of *otx2* and *vox* are partially overlapping in the borders of the shield area (Fig. S1). Same as in zebrafish (Rhinn et al. 2003), the expression of medaka *gbx1* was observed earlier (Fig. 3b) than *gbx2* which was detected slightly later at the end of gastrulation (Fig. 3c, j). We detected early expression of both *pax2* genes in medaka, in particular *pax2.1* at stage 15 (50% epiboly), followed by *pax2.2* at stage 16 (75% epiboly) in the prospective brain area (Fig. 3d, e).The onset of *en2* was detected during the stage 17 (early neurula) (Fig. 3f, j), whereas *wnt1*, *fgf8, gbx2* and *vox* at stage 16 (75% epiboly) (Fig. 3c, g-i). In regard to *pax2* genes, our observations in medaka are in contrast to zebrafish, where the expression onset of *pax2*.1 is almost simultaneous with *fgf8*, *gbx2* and *wnt1* genes (Lun and Brand 1998; Pfeffer et al. 1998; Reifers et al. 1998; Rhinn et al. 2003). Similarly to medaka (Heimbucher et al. 2007; Ristoratore et al. 1999) the expression of zebrafish *gbx2* and *en2* is initially detectable during the late gastrula and the early neurula stage, respectively (Lun and Brand 1998; Rhinn et al. 2003).
The expression of *otx2*, *gbx1, pax2.1* in medaka is triggered during early gastrulation stages in presumptive midbrain and MHB regions while the expression of *vox* gene is detected in the midbrain area in late gastrula stage. This may suggest that *vox* plays role downstream of these genes in MHB. Based on the expression data of *otx2* and *gbx1/2* in medaka (Fig. 3a, b, c) and zebrafish (Fig. 3j), it is probable that both species evolved a specific switch in gene function of *gbx1* instead of *gbx2* (Rhinn and Brand 2001).
In general, the onset of *otx2*, *gbx1*, *gbx2*, *wnt1*, *fgf8* and *en2* in medaka is comparable with the onset of zebrafish homologs and subtle apparent differences might be due to unequal development between these two species. The most notable difference in expression between medaka and zebrafish is the relative onset of *pax2* gene paralogs. Medaka *pax2.1* was expressed much earlier than zebrafish *pax2.1,* and coincides with the onset of *gbx1* during the establishment phase. Moreover, medaka *pax2.2* early expression suggests its role during early neural plate patterning. In contrast, the expression onset of zebrafish *pax2.2* occurs at 5-somite stage, when the neural plate is already transformed into the neural keel (Lun and Brand 1998).
In summary, the spatio-temporal expression profiles described above indicate both conserved and divergent features in the gene regulatory network underlying fish MHB development.

***vox* interferes with the midbrain-hindbrain boundary gene regulatory network**

To investigate the function of *vox* during the medaka early development, we applied a loss of function approach using translation-blocking morpholino oligonucleotides (MO). To test the efficiency of vox MO, synthetic mRNA from the vox_aug_region-eGFP construct (See Material and Methods, Fig. S2a) was injected into embryos alone (Fig. S2a, a´) or with *vox* MO (Fig. S2b, b´). As expected, eGFP from mRNA injection without *vox* MO was expressed in whole embryo (Fig. S2a´), however, its expression was blocked in embryos that were injected both with mRNA and *vox* MO (Fig. S2b´).
The knockdown of *vox* had no or little dorso-ventral (D/V) patterning defects during body axis formation which is rather obvious from the normal morphology exhibited by the vast majority of *vox*

morphants (Fig. S2c-f′). Accordingly, *vox* morphants had none or low penetrance phenotypes in developing MHB (Fig. S2c-f′). We examined the expression pattern of MHB genes, namely *otx2*, *en2*, *wnt1* and *fgf8* (Fig. S2c-f′). We observed very subtle or no changes in *vox* morphants comparing with control MO. The lack of phenotypic changes in medaka *vox* morphants might be due to another gene, which could rescue the early D/V body axis formation and later the predicted function in MHB, as well. It has been reported for zebrafish, that *vox* knockout, or its inactivation by MO injection has little or no effect on early D/V body axis establishment of the embryo (Imai et al. 2001).

For further investigation of *vox* role during the medaka brain development, we also applied a gain of function approach. Since the ectopic expression of *vox* causes strong ventralization of the embryos during the early development of vertebrates (Kawahara et al. 2000; Kozmikova et al. 2013), we chose an inducible heat shock element (HSE) system (see Materials and Methods, Fig. S3). The HSE system allows ectopic induction of a gene of interest in a stage-dependent manner together with fluorescent marker mCherry, using a bidirectional promoter (Fig. S3a). The ectopic expression of *vox* was induced at stage 16 (75% epiboly) (Fig. S3b). Activation of the *vox* HSE expression construct at this time period significantly reduced the appearance of ventralized phenotype in the very early stages.
In order to directly check for *vox* overexpression, we further analyzed mCherry positive embryos by RNA *in situ* hybridization using the probe against *vox*. This confirmed strong upregulation of *vox* expression in all embryonic tissues (Fig. 4i, in detail Fig. S3c, c′). The overexpression of *vox* during late gastrula stage repressed dorso-anterior structures of the embryo and lead to ventralized phenotype (Fig. S3d, e). Phenotypic loss of MHB region was occasioned by *vox* overexpression (insets Fig. S3d, e).
To investigate the MHB development in the *vox* HSE transgenic line, we examined the expression domains of core MHB genes, namely *otx2*, *gbx2*, *pax2.2*, *en2*, *wnt1* and *fgf8*, at stages 20-21, when the MHB patterning is established (Fig. 4). The induced overexpression of the *vox* gene leads to severe defects in the maintenance of isthmic organizer genes. Medaka *otx2* pattern was altered in the MHB region and in the anterior region after the overexpression of *vox* (Fig. 4c, c′, j, j′). Medaka *gbx2* expression specifically disappeared in the anterior-most hindbrain and in the rhombomere 4 (red arrows in Fig. 4d, d′, k, k′). However, the *gbx2* expression in the diencephalon, otic placode and neural crest cells remained unchanged in the control and *vox* transgenic line (asterisks Fig. 4d, d′, k, k′). *pax2.2* expression was downregulated in the MHB region in the *vox* transgenic line upon induction (Fig. 4e, e′, l, l′). The broad expression area of *en2* observed in control embryos was restricted to only a small stripe in the medaka *vox* transgenic line (Fig. 4f, f′, m, m′). The expression pattern of *wnt1* is similar in the *vox* transgenic line and control embryos, except for the intensity, which seems to be stronger in the transgenic line (Fig. 4g, g′, n, n′). The MHB-specific stripe of *fgf8* expression was exclusively missing in the *vox* transgenic line, while the *fgf8* expression persisted in the anterior neural ridge (Fig. 4h, h′, o, o′) and in the tail bud (not shown). The fact that *fgf8* expression is strongly downregulated by *vox* overexpression only in MHB suggests spatially-restricted repression by *vox* in this area (Fig. 4p, q). The loss of *fgf8* expression in MHB development of medaka mimics the situation in the zebrafish *acerebellar* (*ace*) mutant line, in which the expressed Fgf8 is truncated and unable to activate Fgf signaling, resulting in severe brain defects. These defects are mainly caused by disruption of the maintenance of MHB. In the *ace* mutant, the *otx2* expression domain is expanded (Jaszai et al. 2003), *gbx2* expression is lost (Rhinn et al. 2003), and *pax2.1* and *en2* expression domains are gradually reduced (Reifers et al. 1998).

The mechanistic role of *vox* remains enigmatic. In other contexts vent-like proteins function as transcriptional repressors. We therefore speculate that *vox* might restrict anterior expression border of *fgf8*, either directly by binding to *fgf8* transcriptional regulatory sequences, or indirectly by modulating the expression of *gbx2*. The latter scenario is supported by the fact that *gbx2*, a known positive regulator of *fgf8*, is downregulated in *vox* gain-of-function situation (Fig. 4d, d′, k, k′). Future identification of gene(s) functioning redundantly with *vox* in MHB, for example by double knockdown or knockout strategy, might help to reveal the exact position of *vox* within the gene regulatory network governing MHB maintenance. Overall, we described here the onset of early CNS markers during the medaka development. Our study suggests that *vox* homeobox gene is a novel member of midbrain-hindbrain boundary gene regulatory network and plays a role in the development of the neural tube.

Acknowledgements
We are grateful to Jindra Pohorela, Anna Zitova, Ivana Dobiasovska and Vladimir Soukup for technical support. We are grateful to Thomas Czerny for providing reagents and to Sarka Takacova for manuscript proofreading. This work was supported by the Ministry of Education, Youth and Sports (LO1419).

Fig. 1 Conservation analysis of the *vox* loci in four fish species. (**a)** Schematic representation of genes in the *vox* loci of medaka (*Oryzias latipes*), fugu (*Takifugu rubripes*), stickleback (*Gasterosteus aculeatus*) and zebrafish (*Danio rerio*). (**b)** Alignment of fish vox proteins. A high level of conservation is observed in the homeobox domain (black line). A putative Eh-1 motif is detected in the N-terminus (blue line), in addition to two more conserved blocks (Box1 and Box2).

Fig. 2 Early embryonic expression of *vox*. (**a**) *vox* is first expressed at the stage 11 (late blastula). (**b**) At stage 15 (50% epiboly) *vox* surrounds the shield area. The black arrowhead in b-inset marks the shield position. (**c**) At stage 16 (75% epiboly) *vox* is broadly expressed on the ventral side and in two stripes (white arrowheads) in the prospective MHB region of the embryo. (**d**, **e**) Stage 17 shows *vox* expression in the future midbrain, tail bud and non-neural tissue. (**e, e´**) Double RNA *in situ* hybridization was conducted for *vox* (purple) and *pax2.2*, a marker of MHB (red); red arrowhead marks the expression overlap. (**f**, **g**, **h**, **i**) At later stages (**f-i**) *vox* is expressed in the dorsal part of neural rod in the midbrain region, the very posterior region of the embryo (**f**) and in the eye (**f**, **g**). Embryos are shown in (**a** and **b**-inset) animal pole views, (**b-d**, **e´** and **g**-bottom) dorsal views and (**e**, **f**, **g**-top and **i**) lateral views. (**h**): Dorso-ventral section at the level of the midbrain-hindbrain region (dorsal to top, position of section is indicated by a dashed line in (**g**)). Ey, eye vesicle; MHB, midbrain hindbrain boundary; TB, tailbud.

Fig. 3 Medaka *vox* is expressed in the same manner as MHB markers. RNA i*n situ* hybridization of wild-type embryos between stages 15 (gastrula) and 17 (early neurula). All embryos are shown in dorsal view. Dashed line indicates position of the shield in the medaka embryo (**a-e**, **g-i**) or embryonic body (**f**). (**j**) Comparison of the onset of expression of the different genes associated with midbrain–hindbrain junction in medaka (this study) and zebrafish (modified from Rhinn and Brand (2001)).

Fig. 4 Heat-induced overexpression of *vox* results in perturbed expression of the midbrain hindbrain genes. (**a**-**a´**) Schematic view of the orientation of embryos. Lateral (**b-h**) and dorsal (**b´-h´**) views of control embryos, and lateral (**i-o**) and dorsal (**j´-o´**) views of *vox* transgenic line embryos after heat shock. Anterior is to the right. White arrowheads show the position of MHB. The regulatory network obtained from literature data and the hypothetical role of the *vox* gene in the control (**p**) and *vox* HSE transgenic line (**q**). Note the loss of *gbx2* and *fgf8* (**k**, **k´**, **o**, **o´**, red arrows) by ectopic *vox*. Asterisks mark the unchanged *gbx2* expression domains. Ey, eye vesicle; HB, hindbrain; MB, midbrain. Control and *vox* transgenic line embryos were treated for 2h at 39°C, and fixed 24h after induction. i -inset: *in vivo* imaging of whole body mCherry expression after heat shock.

Fig. S1 *vox* expression is overlapping with *otx2* expression in neural plate border and neural keel. (**a**-**e´**) Double RNA *in situ* hybridization for *vox* (purple) and *otx2* (red). Dorsal views at stages 15 (**a-d**) and 17 (**e´**). Lateral view at stage 17 (**e**). Dashed line indicates position of the shield in the medaka embryo (**a-d**) or embryonic body (**e**, **e´**). Arrowhead marks the edge of *vox* domain.

Fig. S2 Knockdown of medaka *vox* by morpholino (MO). (**a**) Schematic drawing of the mRNA construct vox_aug_region-eGFP (**b**) mRNA construct with *vox* MO. Embryos injected with the synthetic mRNA alone (**a´**) and with vox MO (**b´**). (**c-f´**) RNA *in situ* hybridization for *otx2*, *en2, wnt1* and *fgf8* genes on embryos injected with control MO (**c-f**) and vox morpholino (**c´-f´**). Embryos are shown, anterior to right, in (**c-f´** top) lateral view, and (**c-f´** bottom) dorsal view. Arrowheads mark developing MHB.

Fig. S3 The effects of *vox* overexpression during early development. (**a**) mCherry:HSE:vox construct. (**b**) Experimental strategy used in this work. (**c**, **c´**) overexpression of *vox* gene (**c**) and mCherry marker gene (**c´**) after heat shock. (**d**-**e´**) Heat-treated control and mCherry:HSE:vox transgenic line embryos at stage 16. (**d, d´**) Heat shock does not seem to influence development of control embryos at later developmental stages (stage 30). (**e, e´**) Overexpression of *vox* during gastrulation affects development of mCherry:HSE:vox transgenic line embryos. Inset: MHB in detail.

REFERENCES

Bajoghli B, Aghaallaei N, Heimbucher T, Czerny T (2004) An artificial promoter construct for heat-inducible misexpression during fish embryogenesis Dev Biol 271:416-430 doi:10.1016/j.ydbio.2004.04.006

Canning CA, Lee L, Irving C, Mason I, Jones CM (2007) Sustained interactive Wnt and FGF signaling is required to maintain isthmic identity Dev Biol 305:276-286 doi:10.1016/j.ydbio.2007.02.009

Dworkin S, Jane SM (2013) Novel mechanisms that pattern and shape the midbrain-hindbrain boundary Cell Mol Life Sci 70:3365-3374 doi:10.1007/s00018-012-1240-x

Echelard Y, Epstein DJ, St-Jacques B, Shen L, Mohler J, McMahon JA, McMahon AP (1993) Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity Cell 75:1417-1430

Echevarria D, Vieira C, Gimeno L, Martinez S (2003) Neuroepithelial secondary organizers and cell fate specification in the developing brain Brain Res Brain Res Rev 43:179-191

Fabian P, Kozmikova I, Kozmik Z, Pantzartzi CN (2015) Pax2/5/8 and Pax6 alternative splicing events in basal chordates and vertebrates: a focus on paired box domain Frontiers in genetics 6:228 doi:10.3389/fgene.2015.00228

Gilardelli CN, Pozzoli O, Sordino P, Matassi G, Cotelli F (2004) Functional and hierarchical interactions among zebrafish vox/vent homeobox genes Dev Dyn 230:494-508 doi:10.1002/dvdy.20073

Heimbucher T et al. (2007) Gbx2 and Otx2 interact with the WD40 domain of Groucho/Tle corepressors Mol Cell Biol 27:340-351 doi:10.1128/MCB.00811-06

Houart C, Westerfield M, Wilson SW (1998) A small population of anterior cells patterns the forebrain during zebrafish gastrulation Nature 391:788-792 doi:10.1038/35853

Imai Y, Gates MA, Melby AE, Kimelman D, Schier AF, Talbot WS (2001) The homeobox genes vox and vent are redundant repressors of dorsal fates in zebrafish Development 128:2407-2420

Iwamatsu T (2004) Stages of normal development in the medaka Oryzias latipes Mech Dev 121:605-618 doi:10.1016/j.mod.2004.03.012

Jaszai J, Reifers F, Picker A, Langenberg T, Brand M (2003) Isthmus-to-midbrain transformation in the absence of midbrain-hindbrain organizer activity Development 130:6611-6623 doi:10.1242/dev.00899

Jessell TM, Sanes JR (2000) Development. The decade of the developing brain Curr Opin Neurobiol 10:599-611

Joyner AL, Liu A, Millet S (2000) Otx2, Gbx2 and Fgf8 interact to position and maintain a mid-hindbrain organizer Curr Opin Cell Biol 12:736-741

Kawahara A, Wilm T, Solnica-Krezel L, Dawid IB (2000) Antagonistic role of vega1 and bozozok/dharma homeobox genes in organizer formation Proc Natl Acad Sci U S A 97:12121-12126 doi:10.1073/pnas.97.22.12121

Kozmikova I, Candiani S, Fabian P, Gurska D, Kozmik Z (2013) Essential role of Bmp signaling and its positive feedback loop in the early cell fate evolution of chordates Dev Biol 382:538-554 doi:10.1016/j.ydbio.2013.07.021

Lee KJ, Jessell TM (1999) The specification of dorsal cell fates in the vertebrate central nervous system Annu Rev Neurosci 22:261-294 doi:10.1146/annurev.neuro.22.1.261

Li Y, Allende ML, Finkelstein R, Weinberg ES (1994) Expression of two zebrafish orthodenticle-related genes in the embryonic brain Mechanisms of development 48:229-244

Liu A, Joyner AL (2001) EN and GBX2 play essential roles downstream of FGF8 in patterning the mouse mid/hindbrain region Development 128:181-191

Louis A, Nguyen NT, Muffato M, Roest Crollius H (2015) Genomicus update 2015: KaryoView and MatrixView provide a genome-wide perspective to multispecies comparative genomics Nucleic acids research 43:D682-689 doi:10.1093/nar/gku1112

Lun K, Brand M (1998) A series of no isthmus (noi) alleles of the zebrafish pax2.1 gene reveals multiple signaling events in development of the midbrain-hindbrain boundary Development 125:3049-3062

Melby AE, Beach C, Mullins M, Kimelman D (2000) Patterning the early zebrafish by the opposing actions of bozozok and vox/vent Dev Biol 224:275-285 doi:10.1006/dbio.2000.9780

Nicholas K, Nicholas H, Deerfield D (1997) GeneDoc: Analysis and visualization of genetic variation.

Pfeffer PL, Gerster T, Lun K, Brand M, Busslinger M (1998) Characterization of three novel members of the zebrafish Pax2/5/8 family: dependency of Pax5 and Pax8 expression on the Pax2.1 (noi) function Development 125:3063-3074

Raible F, Brand M (2004) Divide et Impera--the midbrain-hindbrain boundary and its organizer Trends Neurosci 27:727-734 doi:10.1016/j.tins.2004.10.003

Reifers F, Bohli H, Walsh EC, Crossley PH, Stainier DY, Brand M (1998) Fgf8 is mutated in zebrafish acerebellar (ace) mutants and is required for maintenance of midbrain-hindbrain boundary development and somitogenesis Development 125:2381-2395

Rhinn M, Brand M (2001) The midbrain--hindbrain boundary organizer Curr Opin Neurobiol 11:34-42

Rhinn M, Lun K, Ahrendt R, Geffarth M, Brand M (2009) Zebrafish gbx1 refines the midbrain-hindbrain boundary border and mediates the Wnt8 posteriorization signal Neural Dev 4:12 doi:10.1186/1749-8104-4-12

Rhinn M, Lun K, Amores A, Yan YL, Postlethwait JH, Brand M (2003) Cloning, expression and relationship of zebrafish gbx1 and gbx2 genes to Fgf signaling Mech Dev 120:919-936

Rhinn M, Lun K, Luz M, Werner M, Brand M (2005) Positioning of the midbrain-hindbrain boundary organizer through global posteriorization of the neuroectoderm mediated by Wnt8 signaling Development 132:1261-1272 doi:10.1242/dev.01685

Ristoratore F et al. (1999) The midbrain-hindbrain boundary genetic cascade is activated ectopically in the diencephalon in response to the widespread expression of one of its components, the medaka gene Ol-eng2 Development 126:3769-3779

Sievers F et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega Molecular systems biology 7:539 doi:10.1038/msb.2011.75

Thermes V, Grabher C, Ristoratore F, Bourrat F, Choulika A, Wittbrodt J, Joly JS (2002) I-SceI meganuclease mediates highly efficient transgenesis in fish Mech Dev 118:91-98

Wilson SW, Houart C (2004) Early steps in the development of the forebrain Dev Cell 6:167-181

Wurst W, Bally-Cuif L (2001) Neural plate patterning: upstream and downstream of the isthmic organizer Nat Rev Neurosci 2:99-108 doi:10.1038/35053516

Zhao J, Lambert G, Meijer AH, Rosa FM (2013) The transcription factor Vox represses endoderm development by interacting with Casanova and Pou2 Development 140:1090-1099 doi:10.1242/dev.082008

Fig1

Click here to download Figure Fig1.tif

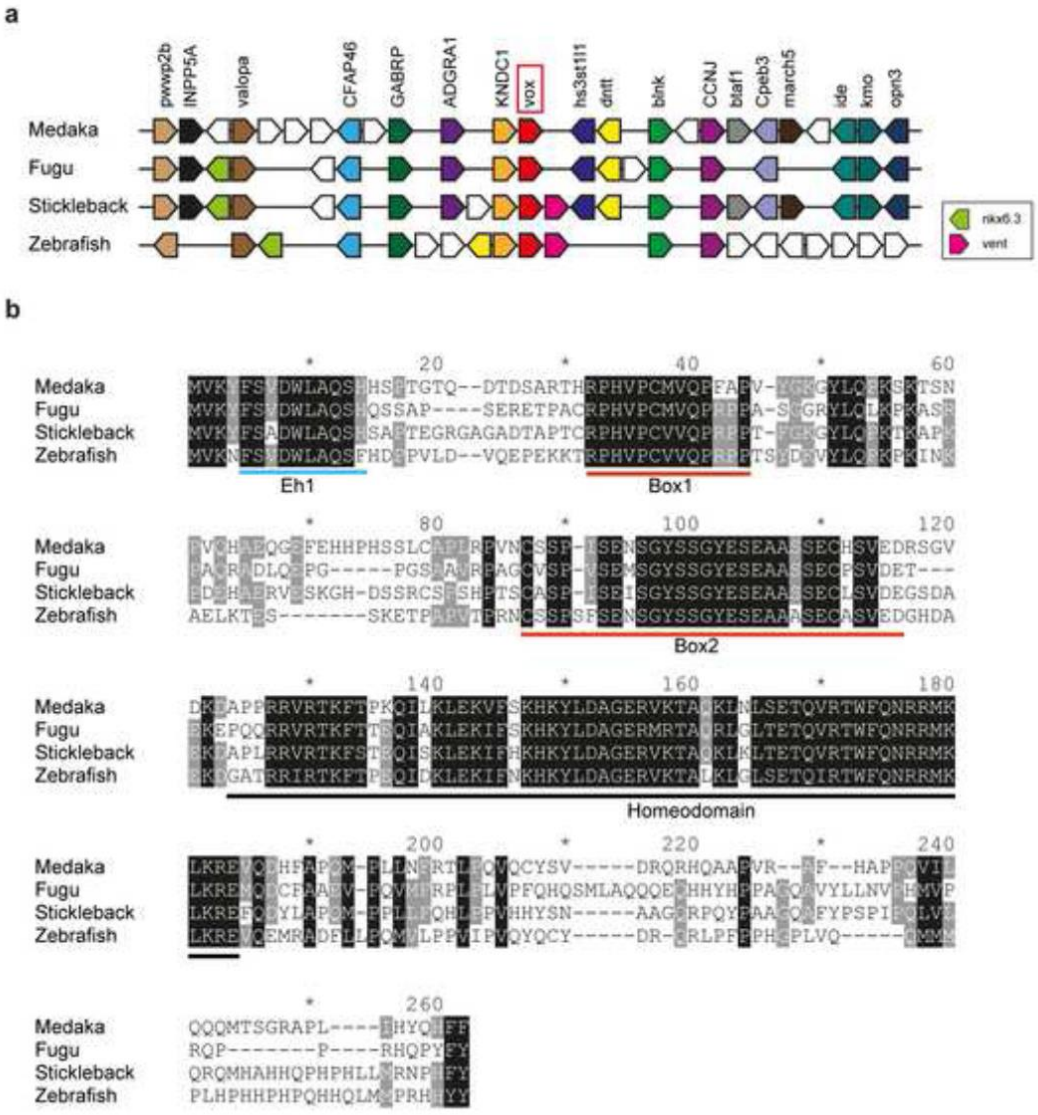


Fig2

Click here to download Figure Fig2.tif

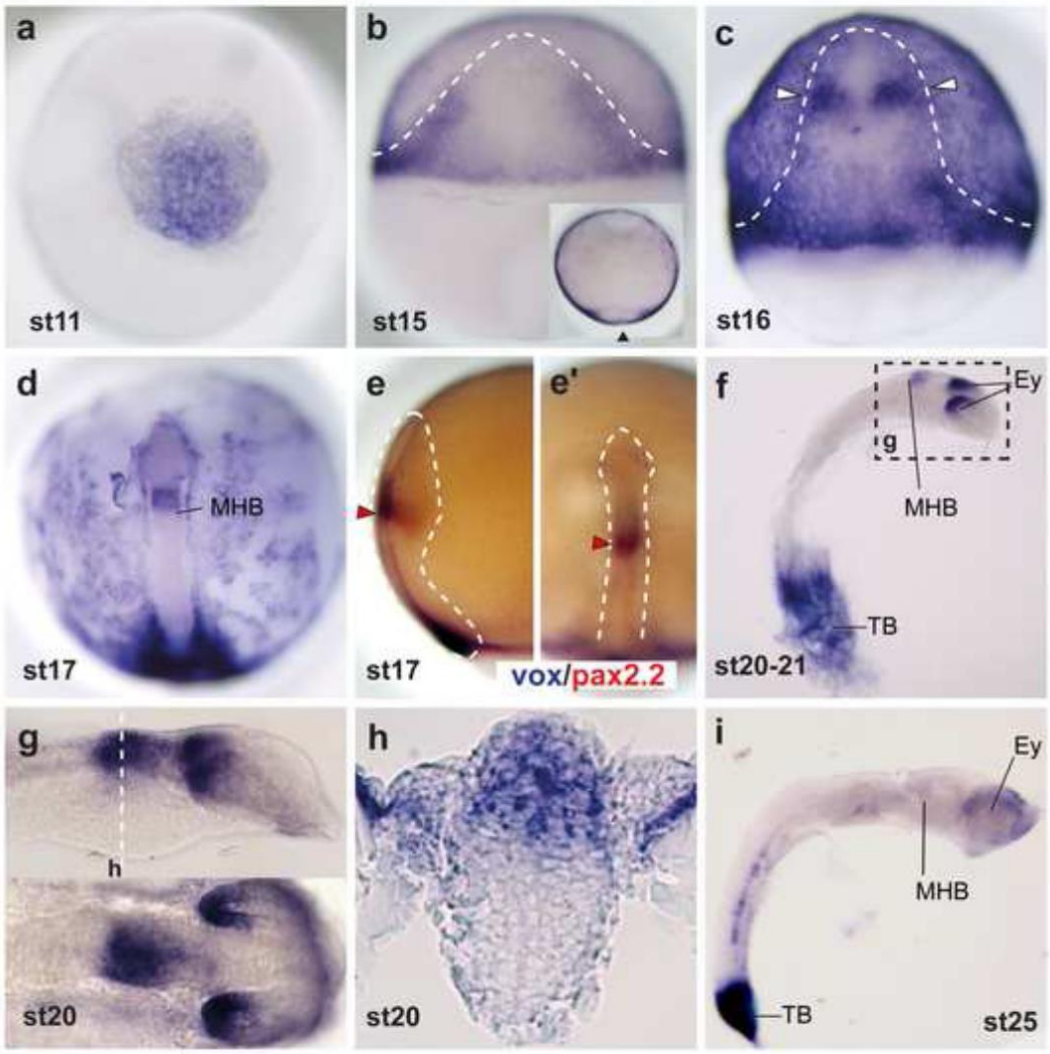


Fig3

Click here to download Figure Fig3.tif

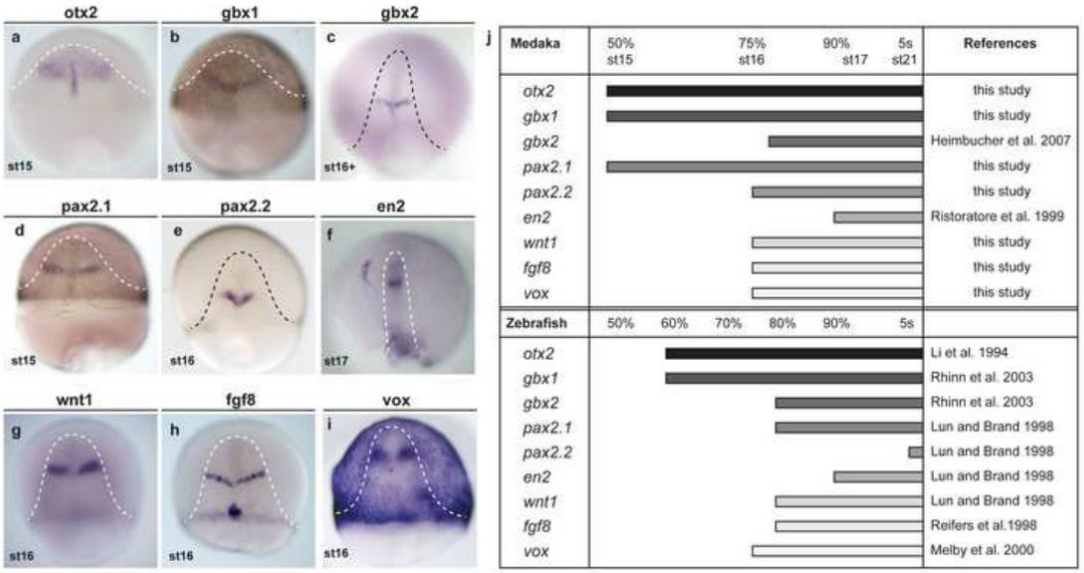
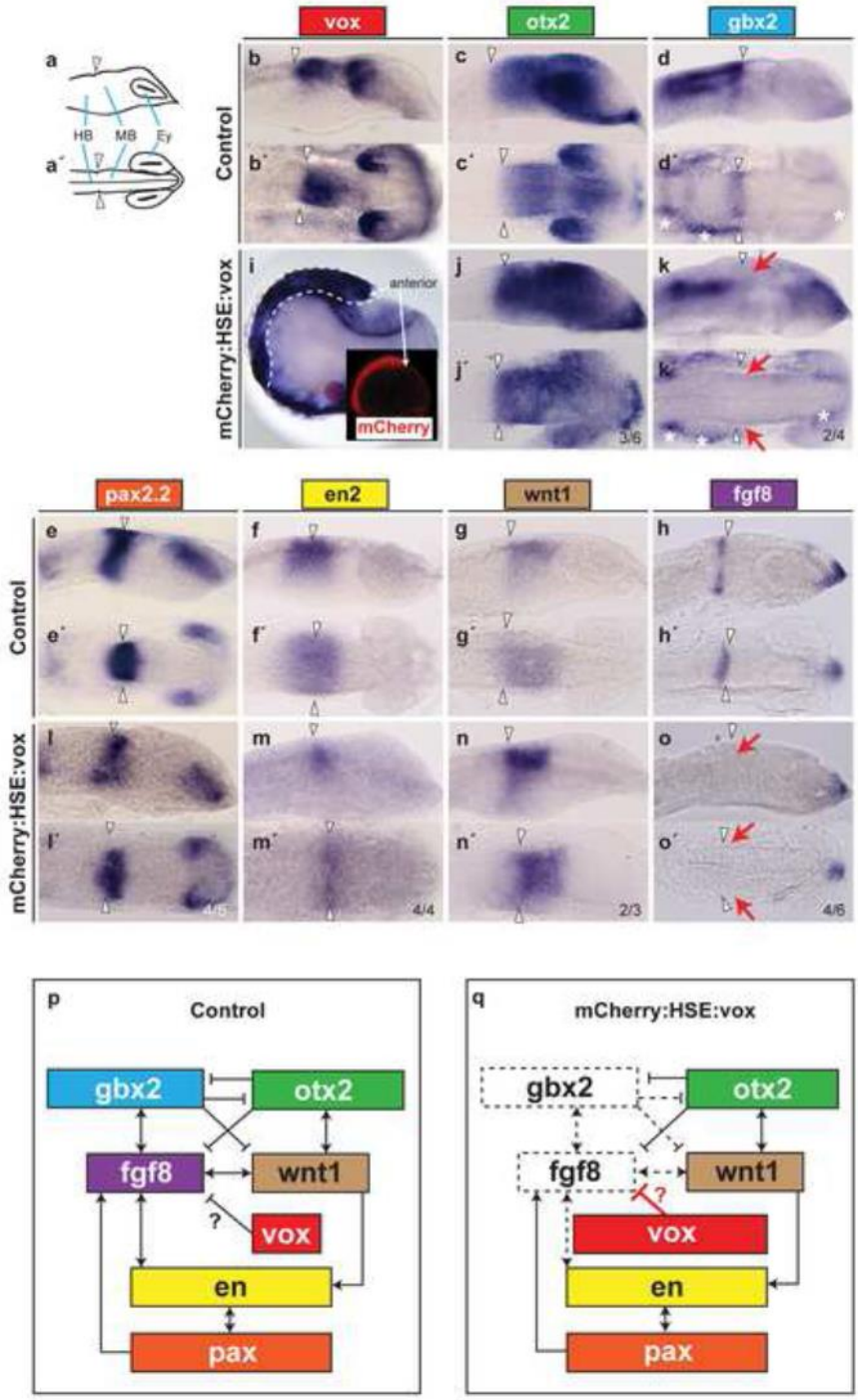
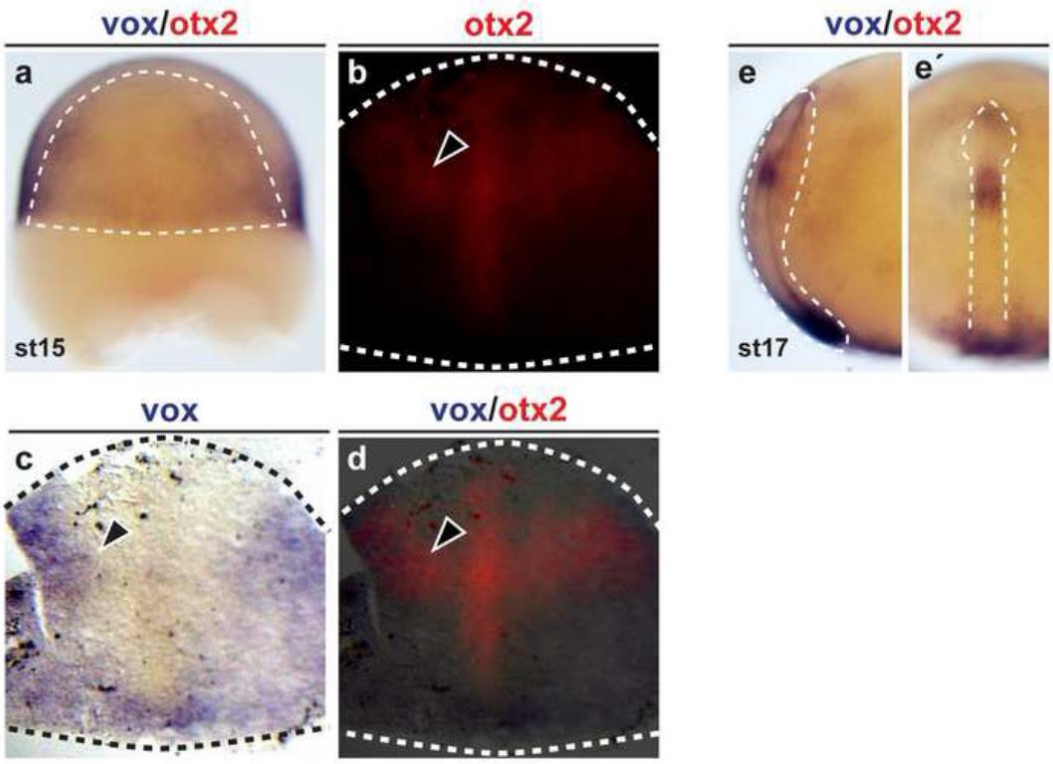


Fig4 Click here to download Figure Fig4.tif

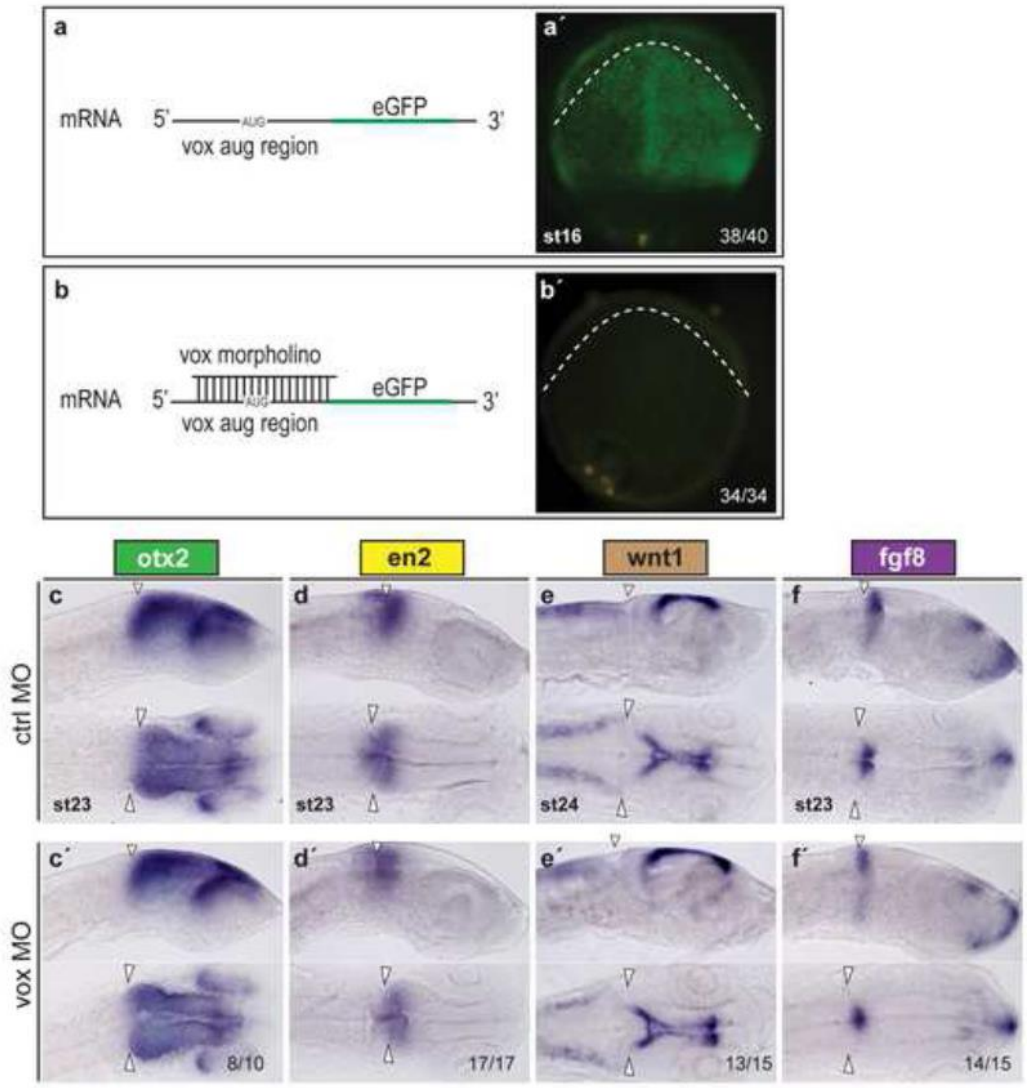


FigS1

Click here to download Figure FigS1.tif

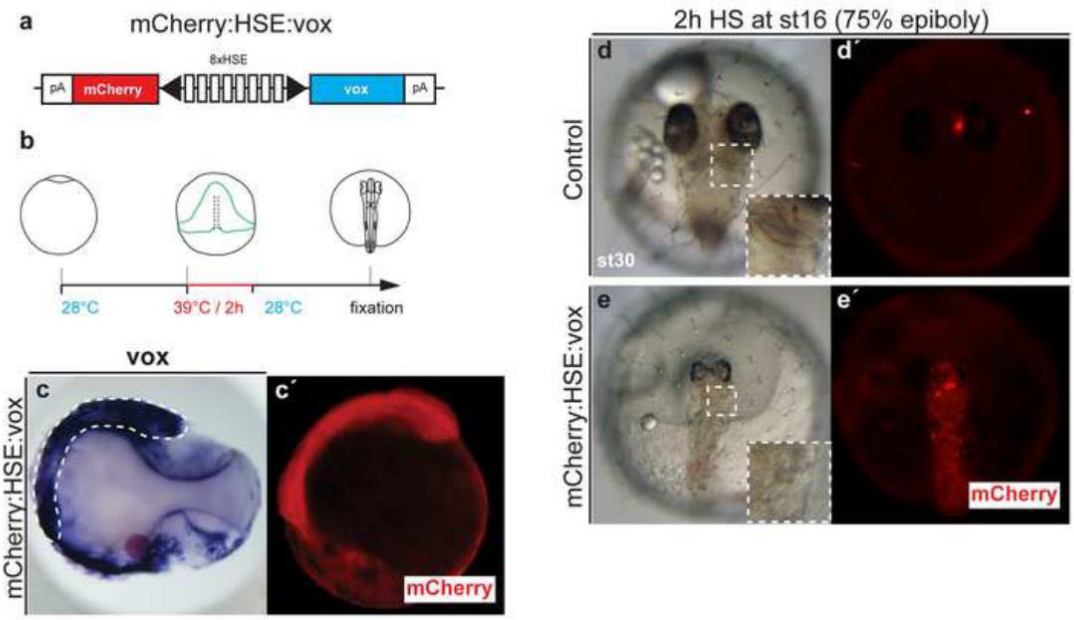


FigS2 Click here to download Figure FigS2.tif



FigS3

Click here to download Figure FigS3.tif



VII.2.3. **Fabian P, Kozmikova I, Kozmik Z, Pantzartzi CN.** *Pax2/5/8* and *Pax6* alternative splicing events in basal chordates and vertebrates: a focus on paired box domain. **Front Genet. 2015 Jul 2;6:22**

frontiers in Genetics

ORIGINAL RESEARCH
published: 02 July 2015
doi: 10.3389/fgene.2015.00228

Pax2/5/8 and *Pax6* alternative splicing events in basal chordates and vertebrates: a focus on paired box domain

*Peter Fabian, Iryna Kozmikova, Zbynek Kozmik and Chrysoula N. Pantzartzi**

Department of Transcriptional Regulation, Institute of Molecular Genetics, Prague, Czech Republic

Paired box transcription factors play important role in development and tissue morphogenesis. The number of *Pax* homologs varies among species studied so far, due to genome and gene duplications that have affected PAX family to a great extent. Based on sequence similarity and functional domains, four Pax classes have been identified in chordates, namely Pax1/9, Pax2/5/8, Pax3/7, and Pax4/6. Numerous splicing events have been reported mainly for *Pax2/5/8* and *Pax6* genes. Of significant interest are those events that lead to Pax proteins with presumed novel properties, such as altered DNA-binding or transcriptional activity. In the current study, a thorough analysis of *Pax2/5/8* splicing events from cephalochordates and vertebrates was performed. We focused more on *Pax2/5/8* and *Pax6* splicing events in which the paired domain is involved. Three new splicing events were identified in *Oryzias latipes*, one of which seems to be conserved in Acanthomorphata. Using representatives from deuterostome and protostome phyla, a comparative analysis of the *Pax6* exon-intron structure of the paired domain was performed, during an attempt to estimate the time of appearance of the *Pax6(5a)* mRNA isoform. As shown in our analysis, this splicing event is characteristic of Gnathostomata and is absent in the other chordate subphyla. Moreover, expression pattern of alternative spliced variants was compared between cephalochordates and fish species. In summary, our data indicate expansion of alternative mRNA variants in paired box region of *Pax2/5/8* and *Pax6* genes during the course of vertebrate evolution.

Keywords: Pax258, Pax6, alternative splicing, paired domain, splice variants

OPEN ACCESS

Edited by:
Hector Escriva, Centre National de la Recherche Scientifique, France

Reviewed by:
Manuel Irimia, Centre for Genomic Regulation, Spain
Simona Candiani, University of Genoa, Italy

****Correspondence:***
Chrysoula N. Pantzartzi, Department of Transcriptional Regulation, Institute of Molecular Genetics, Videnska 1083, Prague 14220, Czech Republic
chrysoula.pantzartzi@img.cas.cz

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

Received: *28 May 2015*
Accepted: *15 June 2015*
Published: *02 July 2015*

Citation:
Fabian P, Kozmikova I, Kozmik Z and Pantzartzi CN (2015) Pax2/5/8 and Pax6 alternative splicing events in basal chordates and vertebrates: a focus on paired box domain. Front. Genet. 6:228. doi: 10.3389/fgene.2015.00228

Introduction

Transcription factors encoded by genes of the paired box (PAX) family are highly conserved throughout metazoan phyla and hold a vital role in embryonic development. The association of different PAX subfamilies with organ and tissue morphogenesis, such as the thymus, central nervous system (CNS), enteric nervous system, kidneys, ear, thyroid, neural crest, vertebrae, and midbrain-hindbrain boundary (MHB) formation has been the object of various studies (summarized in Noll, 1993; Chi and Epstein, 2002; Paixao-Cortes et al., 2013; Blake and Ziman, 2014). Certain members of the PAX family are characterized as the master control genes for eye morphogenesis (Gehring, 1996, 2002, 2012; Kozmik, 2008; Klimova and Kozmik, 2014) and genetic defects are linked to the onset of eye-related diseases, e.g., small eye in mouse or aniridia in human.

Mutations in *Pax* genes are also correlated with diseases of the kidney and the CNS, but various types of cancer as well (see Chi and Epstein, 2002; Lang et al., 2007; Paixao-Cortes et al., 2013; Blake and Ziman, 2014 and references therein).

Gene and genome duplications, followed by gene losses, helped shape PAX gene family, leading to a varying number of *Pax* homologs in metazoan phyla studied so far (reviewed in Noll, 1993; Breitling and Gerber, 2000; Hoshiyama et al., 2007; Paixao-Cortes et al., 2013). Different subfamilies have been identified and are classified according to the similarity in sequence, functional domains they possess as well as their expression patterns (see Stuart et al., 1994; Blake and Ziman, 2014 and references therein). *PaxB* is apparently the oldest member and is present in sponges and cnidarians (Kozmik et al., 2003; Hoshiyama et al., 2007; Hill et al., 2010). *Pox neuro* and single genes from Pax1/9, Pax2/5/8, Pax4/6, and Pax3/7 classes are present in cephalochordates (Short and Holland, 2008; Takatori et al., 2008). *Pox neuro* is present in *Drosophila* (Bopp et al., 1989), but is lost in the lineages of tunicates and vertebrates. Two *Pax258* genes are present in urochordates, due to a duplication prior to ascidian and larvacean diversification (Wada et al., 2003; Canestro et al., 2005). As a result of two rounds of whole-genome duplication (Escriva et al., 2002; Putnam et al., 2008) and subsequent gene losses, the coelacanth *Latimeria chalumnae* possesses nine *Pax* genes, i.e., discrete gene copies from each of the Pax1/9, Pax2/5/8, Pax3/7, and Pax4/6 classes, suggesting that all members of the PAX family were present in the ancestor that gave rise to the tetrapod lineage (Paixao-Cortes et al., 2013). More than nine genes are present in teleost fishes (reviewed in Ravi et al., 2013), due to the so-called third round of genome duplication (Jaillon et al., 2004; Van De Peer, 2004).

All Pax proteins contain a DNA-binding domain in their N-terminus, known as the paired domain (PD), as well as transactivation and inhibitory domains in their C-terminus. The PD consists of 128 aminoacids and is made up by two helix-turn-helix (HTH) subdomains, known as PAI and RED or N-terminal and C-terminal respectively, joined through a linker (Czerny et al., 1993; Xu et al., 1999). The Pax4/6 class contains an additional DNA-binding homeodomain, while the Pax2/5/8 possesses a partial homeodomain and an octapeptide motif, the latter known to interact with members of the Groucho family of co-repressors (Eberhard et al., 2000; Kreslova et al., 2002). Classes Pax1/9 and Pax3/7 both contain the octapeptide motif, yet the former lacks the homeodomain (Chi and Epstein, 2002). *Pax* loci which encode for proteins with truncated paired domain have been identified in *Drosophila*, *C. elegans*, as well as representatives of Hemichordata and Echinodermata (Chisholm and Horvitz, 1995; Cinar and Chisholm, 2004; Howard-Ashby et al., 2006; Friedrich and Caravas, 2011; Ravi et al., 2013).

Gene/genome duplication is a driving force for evolution (Bergthorsson et al., 2007; Maere and Van De Peer, 2010) and this could be nicely exemplified by the well-studied PAX gene family, where many duplicates were preserved in the genome and obtained new functions and new domains of expression (neofunctionalization), or original gene functions were partitioned (subfunctionalization) between duplicates (Pfeffer et al., 1998; Bassham et al., 2008; Kleinjan et al., 2008, reviewed in Holland and Short, 2010).

At posttranscriptional level, alternative splicing is also known to promote evolution, protein diversity and development of novel functions in eukaryotic genomes (reviewed in Nilsen and Graveley, 2010; Chen et al., 2012; Kelemen et al., 2013). In fact, it has been suggested that in the case of *Pax* genes the impact of alternative splicing on functional motifs is more intense than gene duplication and subsequent divergence of the duplicated genes (Short and Holland, 2008). It has been shown that alternative spicing usually takes place in a tissue or developmental stage-specific manner (Wang et al., 2008a; Kelemen et al., 2013). Depending on which exonic segments are cut-out and whether intronic regions are retained in the transcripts, splicing events can be clustered into four major groups, namely (1) exon skipping, (2) alternative 3′-, (3) alternative 5′-splice sites, and (4) intron-inclusion (Koralewski and Krutovsky, 2011). These four types of events can occur independently or in combination with other incidents, such as mutually exclusive exons, alternative initiation and alternative polyadenylation (Wang et al., 2008a; Koralewski and Krutovsky, 2011; Kelemen et al., 2013).

Alternative splicing of *Pax* genes has been observed in various species, from protostomes (Fu and Noll, 1997; Cinar and Chisholm, 2004) to cephalochordates (Glardon et al., 1998; Kozmik et al., 1999; Short and Holland, 2008; Holland and Short, 2010; Short et al., 2012) and vertebrates (Kozmik et al., 1993, 1997; Poleev et al., 1995; Heller and Brandli, 1997, 1999; Lun and Brand, 1998; Short et al., 2012), where multiple incidents from all major groups of splicing events were present. In principle, splice forms seem to have diverged between lineages, some of them are species- or genus-specific (Heller and Brandli, 1999; Short et al., 2012), nevertheless several splice isoforms seem to be evolutionary conserved (Kwak et al., 2006; Short and Holland, 2008; Short et al., 2012; Ravi et al., 2013). The majority of reported splice events regards the transactivation and inhibitory domain in the C-terminal part of the Pax proteins (Kozmik et al., 1993; Ward et al., 1994; Nornes et al., 1996; Tavassoli et al., 1997; Kreslova et al., 2002; Robichaud et al., 2004), still there is an increasing number of events affecting paired domain and consequently DNA binding capacity (Kozmik et al., 1993, 1997; Zwollo et al., 1997; Short and Holland, 2008; Short et al., 2012). One such example is the *Pax6(5a)* isoform, where inclusion of exon 5a (Walther and Gruss, 1991; Glaser et al., 1992; Puschel et al., 1992) leads to a protein with interrupted paired domain that recognizes an altered DNA binding sequence (Epstein et al., 1994). Apparently, this event is quite conserved among vertebrate lineages, with differences in size and peptide sequence of exon 5a between fish and tetrapods (Ravi et al., 2013). In some cases, alternatively spliced *Pax* isoforms exhibit temporally and spatially differentiated expression patterns (Kozmik et al., 1993, 1997; Heller and Brandli, 1997; Short and Holland, 2008) and have been associated with cancer and genetic disorders (reviewed in Wang et al., 2008b; Holland and Short, 2010).

Previous studies have shown that any insertion in the conserved paired domain, no matter if it is a single-aminoacid extension or a whole exon cassette, modifies DNA binding

capacity and attributes differentiated functions to the isoforms bearing the insertion (Kozmik et al., 1997; Azuma et al., 2005).

In the present study, we sought to identify splicing events in Pax2/5/8 and Pax6 classes that affect the paired domain and study the expression patterns of these alternative spliced transcripts. We identified three new splicing events in *Oryzias latipes Pax2* genes, one of which seems to be highly conserved in Acanthomorphata. We detected a re-occurring splicing event in *O. latipes and Danio rerio Pax6* genes generating the exon 5a insertion. Using our data set we tried to elucidate the time point at which the exon 5a-insertion appeared and the extent of its conservation in various phylogenetic groups.

Materials and Methods

Data Collection and *de novo* Gene Annotation

Nucleotide and aminoacid sequences for annotated *Pax2/5/8* and *Pax6* genes were obtained using proper keywords, through NCBI GenBank (Benson et al., 2013), ENSEMBL release 78 (Cunningham et al., 2015), the UCSC Genome Browser database (Karolchik et al., 2014), the SpBase (Sea Urchin Genome Database, Cameron et al., 2009) and the JGI (Grigoriev et al., 2012). The retrieved *Pax* genes were crosschecked using GENSCAN (Burge and Karlin, 1997), BLASTx and version 0.9 of the NNSPLICE splice predictor (Reese et al., 1997).

For various taxonomic groups (e.g., Chondricthyes: Holopocephali) there are available genomes, but no *Pax* genes are annotated in public databases. In order to include representatives from these groups, we conducted BLAST searches against the NCBI GenBank and wgs subdivision (Trace archive), using known homologs from Deuterostomia species. Where required, small contigs or scaffolds were fused using Merger of the EMBOSS software suite (Rice et al., 2000) and gene structure was defined using GENSCAN (Burge and Karlin, 1997), BLASTx and splice predictor (Reese et al., 1997). ScanProsite (De Castro et al., 2006) was used to detect conserved functional domains in newly identified genes. In addition, adjacent genes of the *de novo* predicted *Pax* genes were also predicted/annotated and gene order was compared to known *Pax* syntenic regions through the Genomicus website v78.01 (Louis et al., 2013). In all cases, the NNSPLICE was used for the prediction of possible alternative acceptor and donator sites.

PipMaker (Schwartz et al., 2000) was used along with BLAST, in order to locate putative sequence conservation among species. Alignment of Pax6 paired domains from various species was performed using the MUSCLE algorithm (Edgar, 2004) included in Mega version 5.0 (Tamura et al., 2011).

Expressed Sequenced Tags (ESTs) Retrieval and Analysis

In order to validate already annotated or newly predicted *Pax* homologs, BLAST searches were performed against the ESTs subdivision. Collected ESTs were aligned with predicted coding sequences from genome analyses and mRNA sequences—if available—in order to detect putative splicing events not recognized so far.

Animal Collection

Specimens of *Branchiostoma floridae* were collected from Old Tampa Bay, Florida, USA. Gametes were obtained and embryos were raised, as previously described (Holland and Yu, 2004). *B. lanceolatum* adults were collected in Banyuls-sur-Mer, France, prior to summer breeding season and raised in the lab until spawning. The spawning of males and females was induced by temperature shift (Fuentes et al., 2007). *B. lanceolatum* and *B. floridae* embryos were developed at 16°C and 26°C, respectively. Embryos of inbred strains of *Oryzias latipes* (Cab) and *Danio rerio* (AB) were used for all experiments. *O. latipes* and *D. rerio* embryonic stages were determined according to Iwamatsu (2004) and Kimmel et al. (1995). Housing of animals and *in vivo* experiments were performed after approval by the Animal Care Committee of the Institute of Molecular Genetics (study ID#36/2007) and in compliance with national and institutional guidelines (ID#12135/2010-17210).

RNA Isolation and Reverse Transcription

Total RNA was isolated from embryos using the Trizol reagent (Ambion). Random-primed cDNA was prepared in a 20 µl reaction from 500 ng of total RNA using SuperScript VILO cDNA Synthesis kit (Invitrogen).

Screen for Alternative Splicing and RT-PCR Analysis

cDNA was subjected to PCR using DreamTaq polymerase (Thermo Scientific) for 30 cycles under the following conditions: 1 min at 98°C, 30 s at 60°C, 30 s at 72°C. Primers for this analysis are provided in **Table 1**. PCR products were analyzed on 2.5% agarose gel and bands of interest were eluted, cloned to pCR-Blunt II (Invitrogen) and sequenced (GATC Biotech sequencing service, Germany).

Results

Pax2/5/8 Splicing Events in Chordates

Exhaustive search through databases and literature, in combination with *de novo* analysis of available ESTs and mRNA sequences (Table S1), revealed numerous splicing events in chordate members of the *Pax2/5/8* class (**Figure 1**). Some of these events seem to characterize specific orthologs and are present in cephalochordates, fish and mammals (e.g., exon 2 of *Pax5* gene), while others are much less conserved (e.g., exon 3a of mouse *Pax5*). *Branchiostoma floridae* appears to experience the largest number of splicing events, however no event of insertion in the paired domain has been reported so far and no such event could be predicted using splice prediction software or available ESTs/mRNA sequences.

A single *Oryzias latipes Pax2* gene, namely *OlPax2.2*, located on chromosome 19 (NC_019877), has been used in previous studies (Paixao-Cortes et al., 2013). Our search through NCBI revealed an annotated Pax-2a-like gene on chromosome 15 (NW_004088010.1). It must be noted that the aminoacid sequence encoded by the first half of this gene exhibits no similarity to the paired domain of other *Pax* genes and it is not supported by ESTs, a fact that indicates an erroneous gene

TABLE 1 | Summary of primers used in RT-PCR reactions in the present study.

Species	Gene	Sequence 5′->3′
Branchiostoma	Pax258	F: AATGGGTGCGTGTCCAAGATT
		R: AACGCTGGGGATGTTCTCATT
	Pax46	F: GTCCCACGGCTGTGTCAGTAA
		R: TCGTTGTCACAAATGCCTTCC
Oryzias latipes	Pax2.1	F: GCAGCGGATCGTGGAGCTT
		R: GAACATAGTGGGGTTTTGGCGC
		R*: CTCTGAAATGCCTTCTGATAG
	Pax2.2	F: GCAGCGGATTGTGGAGCTG
		R: AAACATTGTTGGGTTTTGTCTT
	Pax6.1	F: CCACCAGGCAGAAAATAGTGGAACTT
		R: ATCTTGCTCACGCAGCCGTTGGAT
	Pax6.3	F: AACCAGCTCGGGGGGGTATTTGTG
		R: CTCCAATGGCCCGGGGACGG
Danio rerio	Pax2.1	F: ACCAGCTAGGAGGGGTGTTT
		R: CCAGGCGAACATTGTAGGAT
	Pax2.2	F: CGACAGCTGAGGGTCAGTC
		R: CGAACATGGTGGGATTTTGT
	Pax6.1a	F: CCCGACTCCACGAGACAGAAAATAGTT
		R: ACCCAAGATTTTACTCACGCAGCCGTTG
	Pax6.1b	F: CCGGACTCCACCAGACAGAAGATCGTC
		R: CCAGAATCTTGCTCACGCACCCATTC

For Branchiostoma species the same set of primers was used for each gene class amplified. Oryzias latipes Pax2.1 reverse primer marked with asterisk was used to amplify alternative exon 2a along with Pax2.1 forward primer.

prediction, caused by a non-sequenced area in this genomic region. We assume that the first coding exon as well as the two exons coding for the paired domain of *OlPax2.1* are located in this non-sequenced region. The record of an unplaced scaffold (NW_004093539, Table S1) was retrieved through BLAST. It apparently corresponds partly to the non-sequenced region of chromosome 15 and contains the 5′ UTR, the first exon of *OlPax2.1* gene and part of the first intron.

In order to retrieve more information on *OlPax2.1* and to elucidate how many *OlPax2* (either *Pax2.1* or *Pax2.2*) transcripts exist, we searched for different ESTs and mRNA sequences using *OlPax2.2* and *Danio rerio Pax2.1* as queries. There are only two ESTs (AM320053 and AM321390) for *OlPax2.1* that contain the first coding exon, as well as the complete exon 2 and part of exon 3, encoding for N- and C- paired subdomains, respectively. Using the genomic scaffolds and available ESTs collectively (Table S1), an almost complete *OlPax2.1* gene was reconstructed (**Figure 1**). For *OlPax2.2* gene, one cDNA sequence and four ESTs were retrieved (Table S1), comparison of which revealed both 5′ and 3′ alternatively spliced parts of exon 2 encoding for the N-terminal of paired domain (see **Figure 1**).

Exon-to-exon comparison of *Pax2/5/8* genes between *O. latipes* and *D. rerio*, shows that in principle there is conservation in the sequence, number, size, and borders of exons and some indication for alternatively spliced exons in *OlPax2* genes (light gray boxes in **Figure 1**), which are not suggested by the available ESTs. Retrieved ESTs encoding both *OlPax2.1* and *OlPax2.2* support a 5′ alternative splicing donor site in exon 2 (**Figure 1**, Table S1); the insertion is 12 bp long, and results to four additional aminoacids, exactly at the beginning of the paired domain. The same insertion has been reported for *D. rerio Pax2.1* (Lun and Brand, 1998), as well.

In the present study we identified two splicing events that lead to insertion of extra aminoacids in the paired domain of *OlPax2* genes. More specifically, an alternatively spliced 21-bp exon was detected between exons 2 and 3 of fish *Pax2.1* genes, which is annotated in some species (e.g., *Poecilia reticulata* and *Maylandia zebra*). This exon could not be detected *in silico* in the genome of *O. latipes*, due to the fact that the intron between exons 2 and 3 is not sequenced. Through BLAST searches and *de novo* analysis of *Pax2.1* genes we spotted this exon in numerous representatives from different orders of Acanthomorphata (Table S2), while PipMaker alignment reveals a high degree of sequence conservation among compared species (**Figure 2**). A putative exon with proper splice sites has been identified in the respective genomic region of three Cyprinidae species (*D. rerio*, *Pimephales promelas, and Cyprinus* carpio). Even though this exon is highly conserved in these species, the encoded aminoacids are quite dissimilar from those of the Acanthomorphata 21-bp exon (**Figure 2**). In both cases, inclusion of this exon leads to an alternative transcript, which incorporates seven extra aminoacids toward the end of the a3 helix of the PAI subdomain (**Figure 2**).

In regard to *OlPax2.2*, the available mRNA sequence in GenBank and our analysis revealed a 24-bp in-frame extension at the 3′ end of exon 2 (**Figures 1**, **3**, Table S1), which does not alter the downstream translation (**Figure 3**). This isoform, to which we will refer as *OlPax2.2(ext24+)*, is due to an alternative splicing donor downstream the canonical one (**Figure 3**). In *D. rerio*, a similar isoform is neither supported by splicing prediction software nor by available mRNA sequences.

It should be noted that splicing prediction analysis of *OlPax2.1* gene revealed the presence of an alternative donor site in exon 2, upstream the canonical one. Deletion of 35 bp at the 3′ end of exon 2 causes a frameshift and leads consequently to a premature stop codon at the beginning of exon 3 (Figure S1). A similar donor site was not *in silico* identified in *D. rerio* (data not shown).

Comparative Analysis of 5a-exon Insertion in Pax6 Paired Domain

In teleosts and tetrapods studied so far, the major part of Pax6 paired domain is encoded by two exons, responsible for the N- and C-subdomains, with a size of 131 and 216 bp, respectively. In Tetrapoda and in the *Pax6.1* copy of teleosts, a small exon of varying size (36–42 bp), namely 5a, has been shown to be included in alternative transcripts, causing an in-frame insertion in the paired domain (Ravi et al., 2013). We wanted to identify at which point of evolutionary history this exon appeared and investigate a putative correlation of the appearance of this exon with the exon/intron organization of *Pax6* homologs. For this reason, already annotated *Pax6* homologs were collected from public databases and available genomes and EST sequences from non-jawed vertebrates,

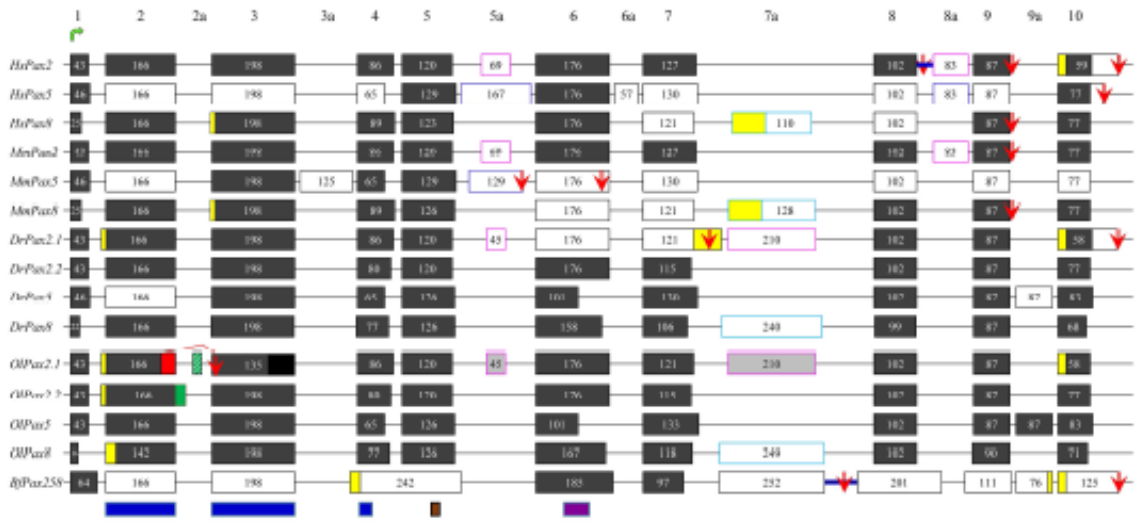


FIGURE 1 | Exon-intron organization and alternative splicing events of *Pax2/5/8* representatives from vertebrates and cephalochordates. Dark gray and white boxes represent constitutive and alternatively spliced exons, respectively. Yellow boxes denote alternatively spliced parts of exons, due to different 5′ or 3′ splicing donors/acceptors. The suffix "a" is used for non-canonical exons, characteristic of the different *Pax2/5/8* genes (different colors of outline is used for different orthologs). For *O. latipes*, light gray boxes were predicted due to high similarity to respective exons in *D. rerio*, whereas black box shows the non-sequenced part of exon 3. Lines represent introns (not drawn to scale). Blue thick lines represent intron retention events. Green arrow points to exon containing translation initiation, while red arrows stand for alternative stop codons. Blue, brown, and purple boxes define borders of paired domain, the octapeptide, and partial homeodomain, respectively. Alternative splicing events detected in the present study are indicated by red and green boxes (alternative 5′ splice donors) and green stripped box (exon cassette). Previously published data are included (Dressler et al., 1990; Krauss et al., 1991; Kozmik et al., 1993, 1997, 1999; Ward et al., 1994; Poleev et al., 1995; Zwollo et al., 1997; Lun and Brand, 1998; Pfeffer et al., 1998; Borson et al., 2002; Robichaud et al., 2004; Kwak et al., 2006; Short and Holland, 2008; Arseneau et al., 2009; Busse et al., 2009).

cephalochordates, tunicates, hemichordates, echinoderms, as well as *Drosophila* and *C. elegans* were analyzed (Table S3).

Analysis of a genomic scaffold from *Leucoraja erinacea*, containing the *Pax6* ortholog, provides evidence that besides Holocephali (Ravi et al., 2013), the 5a exon is also present in Elasmonbranchii, the second subclass of Chondricthyes. Unfortunately, no genomes from hagfishes are publicly available, yet the two *Pax6* mRNA sequences that were retrieved from *Eptatretus bergeri* (Table S3), do not provide any indication of an exon5a-like insertion in the paired domain.

In regard to Hyperoartia, genomic scaffolds containing parts of the *Pax6* genes from *Petromyzon marinus* and *Lethenteron japonicum* were retrieved and analyzed, along with two *Pax6* mRNA from the species *L. japonicum* and *Lampetra fluviatilis*. Apparently, there are more than one *Pax6* genes in the *L. japonicum* genome, yet the low genome coverage for both *P. marinus* and *L. japonicum* (5× and 20×, respectively) does not allow for safer conclusions. In all cases, there is no evidence for the existence of the exon 5a in Hyperoartia.

Existing models for *Stongylocentrotus purpuratus* predict two truncated neighboring Pax6 proteins that contain either the paired or the homeobox domain. Taking into account the provided information in SpBase mentioning that the gene models are incomplete and an intervening sequence appears to be missing in a scaffold gap in the Spur3.1 assembly (Howard-Ashby et al., 2006; Cameron et al., 2009), we re-evaluated the prediction and tried to re-construct the *SpPax6* homolog.

Our focus was on the exon-intron structure in the region of paired domain (**Figure 4**). It is apparent, from our analysis, that the size and borders of the two exons encoding the main part of paired domain underwent various changes in different taxonomic groups (**Figure 4**). More specifically, in basal deuterostomes, such as Echinodermata and Hemichordata, there is one large exon with a size of 347 bp encoding for the first 115 aminoacids of the paired domain. Therefore, there is no intervening non-coding region in the respective position to vertebrate's intron, or in other words "space" for insertion of an alternatively spliced exon in the paired domain. In cephalochordates and tunicates, this large exon has split into two exons, first of which has a size of 166 bp, still larger than the N-subdomain encoding exon in vertebrates, and the second one is 181 bp long, slightly smaller than the respective exon in vertebrates. The 166 bp exon encodes a peptide including the first four aminoacids of a3 helix of the N-subdomain (Xu et al., 1995) and ends shortly after the position where the exon 5a is inserted, i.e., downstream of a3 helix. Thorough *in silico* search in the intronic sequence flanked by the paired-encoding exons in tunicates and cephalochordates did not reveal a putative alternatively spliced exon cassette similar to the exon 5a (**Figure 4**, Figure S2). It seems that paired-encoding

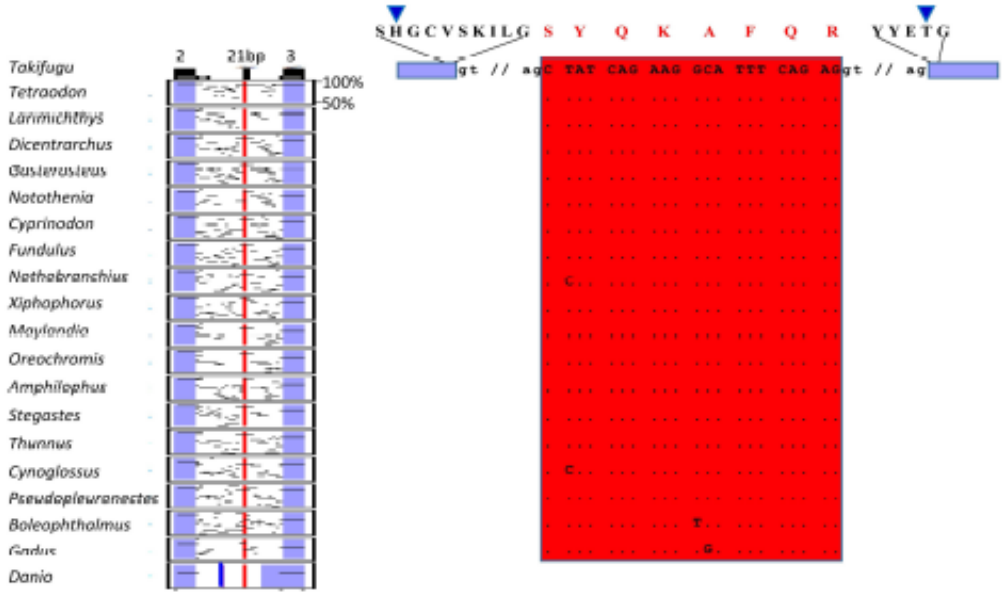


FIGURE 2 | Alternatively spliced isoform *OlPax2.1(2a+)*. Pip diagram of paired-encoding (2 and 3, blue boxes) and alternatively spliced exon (21 bp, red box) in Acanthomorphata species. Blue thick line represents alternatively spliced exon predicted in three Cyprinidae species. Dots stand for conserved residues, small case letters correspond to intronic nucleotides. Blue arrowheads point to PAI a3 helix.

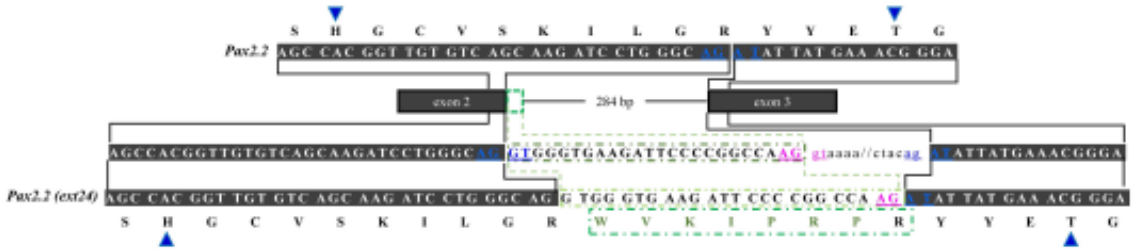


FIGURE 3 | Alternatively spliced isoform *OlPax2.2(ext24+)*. Dark gray boxes represent constitutive exons, green-outlined box shows part of intron 2 (24 bp) retained in *OlPax2.2*. Thin line represents intron 2 and intronic sequences are in small case letters. Canonical splice sites are in blue letters. Blue arrowheads point to PAI a3 helix. Green letters represent the extra aminoacids incorporated in *OlPax2.1(ext24+)*, when the downstream alternative splicing donor (pink letters) is used.

exons obtained fixed borders and size of 131 and 216 bp for PAI and RED, respectively before diversification of cyclostomes and preserved them throughout vertebrates, nonetheless "birth" of exon 5a probably appeared in Gnathostomata (**Figure 4**, Figure S2). It should be noted that the major re-arrangements concern the genomic region encoding for the PAI a3 helix, in contrast to the high conservation observed toward the C-subdomain.

Developmental Expression of PAI-RED Isoforms

To verify *in silico* predicted alternative splice isoforms (**Figures 2–4**, Figure S1) and compare their expression across various developmental stages, we performed RT-PCR using RNA from different embryonic stages of *Branchiostoma lanceolatum*, *B. floridae*, *Oryzias latipes*, and *Danio rerio* (**Figure 5**).

In agreement with our *in silico* analysis, RT-PCR (**Figure 5**) and DNA sequencing (data not shown) using primers located in exons encoding PAI and RED domains revealed that *B. lanceolatum* and *B. floridae* express single *Pax258* and *Pax46* isoforms.

Three splice isoforms of *OlPax2* genes have been *in silico* predicted. Lack of information concerning intron 2 of *OlPax2.1* prohibited the *in silico* detection of the 21-bp exon, characteristic of Acanthomorphata. However, the presence of 21-bp exon in the *O. latipes* genome was experimentally verified, using a proper

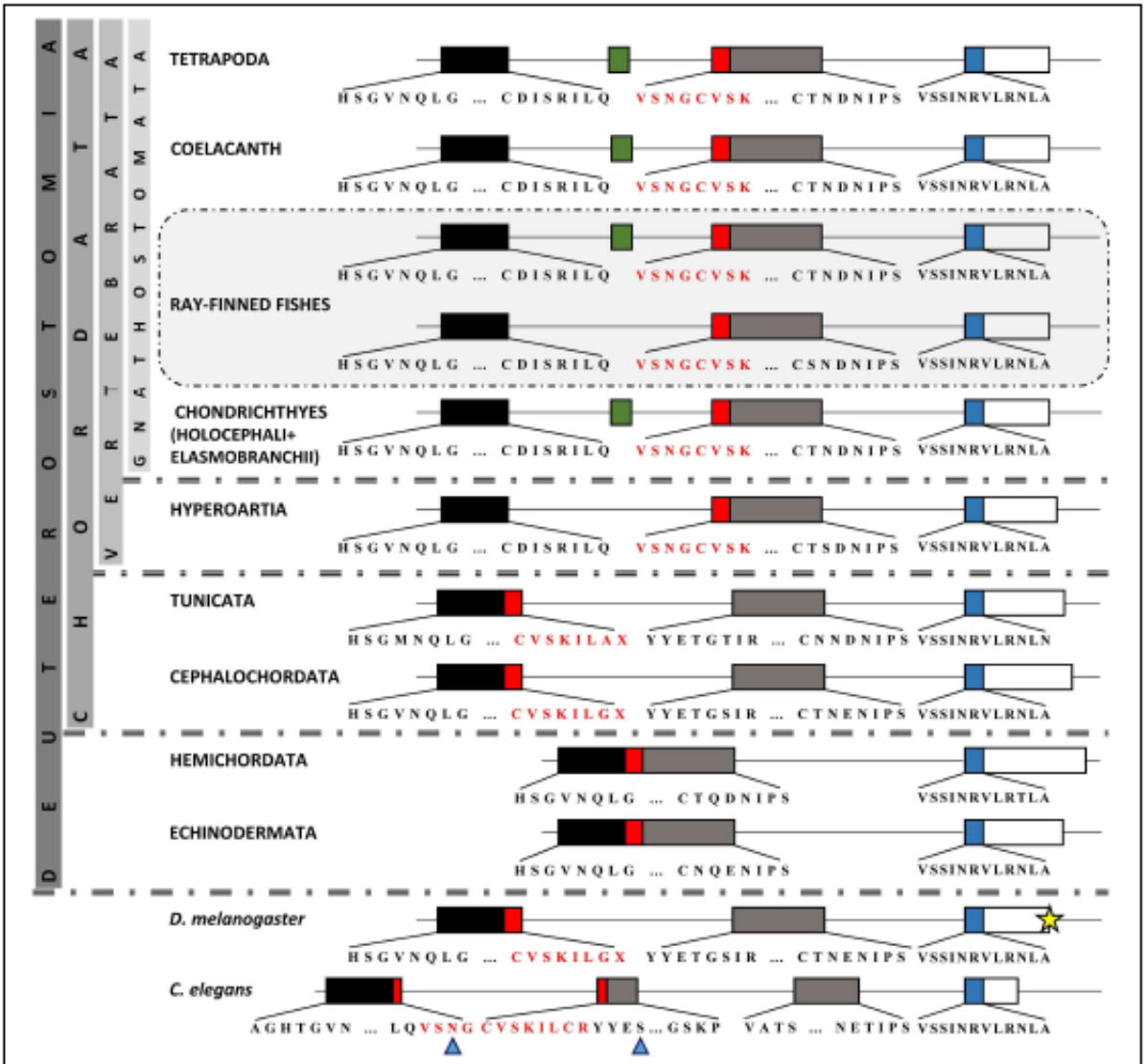


FIGURE 4 | Comparison of Pax6 paired domain among representatives from Protostomia and Deuterostomia. Genomic organization of paired domain-encoding exons. Boxes represent exons, thin lines represent introns (the latter are not drawn to scale). Different colors of exons correspond to different subdomains. Green box represents the previously reported exon 5a (Walther and Gruss, 1991). Red box represents the section of paired domain that underwent re-arrangement throughout evolution. For practical reasons, the third exon (865 bp) in *Drosophila* is not depicted to its full length (yellow star). The residue X reveals an asymmetrical exon. Blue arrowheads point to PAI a3 helix.

set of primers (**Table 1**), one of which was specifically designed on the 21-bp exon characteristic of Acanthomorphata (**Figure 2**). Expression of the alternatively spliced *OlPax2.1(2a+)* isoform was detectable from neurula and later developmental stages (**Figure 5**). In contrast, the truncated isoform *OlPax2.1(del35)*, which results from a deletion in exon 2 and a premature stop codon in the 3rd exon (Figure S1), was present across all developmental stages (**Figure 5**). *OlPax2.1(del35)* was expressed at much lower level than dominant *OlPax2.1(2a-)* isoform. Sequencing of this isoform revealed that it makes use of the alternative donor site in exon 2, as predicted by *in silico* analysis (Figure S1). The extended isoform *OlPax2.2(ext24+)* (**Figure 3**) was present at detectable level throughout the examined stages (**Figure 5**). In the case of *D. rerio Pax2.1* and *Pax2.2* genes, no alternative splice variants were detected, in agreement with splicing prediction analysis.

The alternatively spliced isoform *OlPax6.1(5a-)* was expressed approximately at the same level as isoform *OlPax6.1(5a+)*

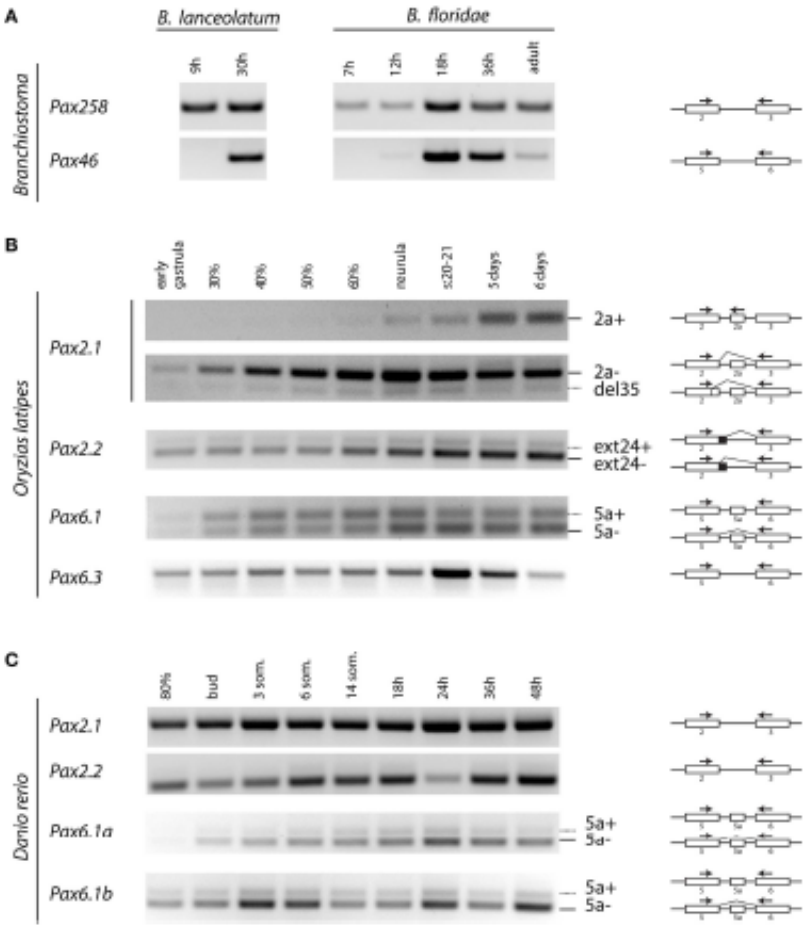


FIGURE 5 | Temporal expression of *Pax2* and *Pax6* genes during development of *Branchiostoma lanceolatum*, *B. floridae*, *Oryzias latipes*, and *Danio rerio* embryogenesis. cDNA sequences were amplified across the paired encoding exons, as indicated by arrows. **(A)** For *B. lanceolatum* and *B. floridae*, single isoforms were detected for *Pax258* and *Pax46* genes. **(B)** For *O. latipes Pax2.1*, an isoform containing exon 2a was detectable from neurula stage. *Pax2.1(del35)* and *Pax2.2(ext24+)* as well as Pax6.1(5a) isoforms are expressed across all developmental stages. Pax6.3 gene does not possess exon 5a. **(C)** No alternatively spliced isoforms were detected in the case of *D. rerio Pax2.1* and *Pax2.2* genes. Isoforms including exon 5a are present for both *Pax6.1a* and *Pax6.1b* genes.

(**Figure 5**). In agreement with previous studies (Ravi et al., 2013), *O. latipes* Pax6.3 gene does not possess the equivalent of exon 5a. Variants bearing exon 5a were observed for both *D. rerio Pax6.1a* and *Pax6.1b* genes, and in both cases, expression level of isoform *5a-* was relatively higher than isoform *5a+*.

In silico and experimental data, collectively, demonstrate increased complexity of splicing events in vertebrate paired domain of *Pax* genes in comparison to cephalochordates.

Discussion

Paired box (*Pax*) genes encode for transcription factors that are considered key players in organogenesis and embryonic development. The presence of *Pax* genes in a variety of organisms and the evolution of the PAX family has been the object of various studies (Hill et al., 2010; Paixao-Cortes et al., 2013; Ravi et al., 2013). Whole-genome duplications as well as lineage-specific gene duplications provide additional possibilities for diversified evolution and/or speciation (Bergthorsson et al., 2007; Maere and Van De Peer, 2010). These processes are considered to have played important role in shaping the number of *Pax* homologs in various taxonomic groups (see Paixao-Cortes et al., 2013; Ravi et al., 2013 and references therein), but same applies for alternative splicing, a posttranslational mechanism that also promoted evolution and complexity of Pax proteins (Glardon et al., 1998; Short et al., 2012).

In the present study, we wanted to evaluate the degree of alternative splicing taking place in various lineages, as well as

to identify splicing events that are either evolutionary conserved or characteristic of cephalochordates and not vertebrates or vice versa. For this purpose, we collected annotated homologs from Pax2/5/8 and Pax4/6 classes from public databases (NCBI, Ensembl, UCSC, JGI and SpBase). Furthermore, we analyzed *de novo* genomes, ESTs and mRNA sequences from different species, in order to enrich our dataset with taxonomic groups not present in previous studies. Our second focus was *Pax* isoforms from cephalochordates and vertebrates, that differ in the paired domain, and their expression patterns across different developmental stages.

Apart from partially reconstructing the *Oryzias latipes Pax2.1*, using available scaffolds and EST sequences, we identified three new splicing events in the *Pax2* genes of *O. latipes*. The *OlPax2.1(2a+)* isoform is reminiscent of the *5a* isoform found in *Pax6* homologs (Ravi et al., 2013, this study), as it incorporates a 21-bp in-frame-exon located in the intron between the two exons encoding for the paired domain. Our analysis showed that this exon is present in numerous species from various orders of Acanthomorphata and exhibits a high degree of conservation among compared species (**Figure 2**). We presume that sequence conservation of this mRNA splice form over a wide phylogenetic distance also implies conservation of this isoform's function. A similar exon in terms of location, yet quite divergent in terms of sequence, was *in silico* predicted only in three Cyprinidae species (*D. rerio*, *Pimephales promelas*, and *Cyprinus carpio*).

The second alternatively spliced isoform, namely *OlPax2.2(ext24+)*, results from the use of an alternative splice donor downstream the canonical one at the end of *OlPax2.2* exon 2. In this case, extra aminoacids are incorporated in the middle of a3 helix of the PAI subdomain, with no influence on the downstream sequence. A similar isoform could not be detected in *D. rerio*, neither experimentally nor *in silico*. The sequences surrounding the normal splice junctions of exon 2-intron 2 are highly conserved between *D. rerio* and *O. latipes Pax2.2* genes, yet there is no proper donor-acceptor site in the region of *D. rerio* (AG-AG) that corresponds to the alternative splice site of *O. latipes*.

Both *OlPax2.1(2a+)* and *OlPax2.2(ext24+)* transcripts bear an insertion in the recognition a3 helix of PAI subdomain. Previous studies on insertions in the paired domain of *Pax* genes have proven that, regardless of the number of the inserted aminoacids, disruption of this helix, which is responsible for all major groove DNA contacts of the N-terminal subdomain (Xu et al., 1995, 1999) is expected to inactivate the DNA-binding function of the N-terminal HTH motif, which subsequently leads to severe restriction in the DNA-binding sequence specificity of the paired domain (Kozmik et al., 1997).

The importance of alternative splicing as a mechanism for divergent evolution is established. In the case of *Pax* genes, the fact that insertions in the paired domain may preferentially guide Pax proteins, namely Pax6(5a) and Pax8(S), to the control region of genes containing a modified binding site (5aCON-like sequence, Kozmik et al., 1997), in other words insertions add new target-genes in the repertoire of genes controlled by *Pax* genes, may be indicative of a mechanism through which alternative splicing contributes to the increase of complexity at the level of protein function.

The isoform *OlPax2.1(del35)* makes use of an alternative 5′ splicing donor, upstream of the normal splicing site in exon 2 (N-terminal of paired domain). As mentioned before, the exact junction sequence between exon 2 and intron 2 is not known, nevertheless, the sequence at the normal end of exon 2 (CAG) is in agreement with the optimal consensus for 5′ splice sites (Stephens and Schneider, 1992), in contrast to the sequences at the alternative upstream 5′ splice donor (CGG/GT, Figure S1). As it has been observed before for the *Pax8* gene (Kozmik et al., 1997), there is a higher abundance and constitutive splicing of the *Pax2* mRNA relative to the alternative transcript (**Figure 5**), a fact that could be attributed to different affinities by which the spliceosomes may recognize the two 5′ donor sites. A similar truncated isoform could not be detected neither during *Pax2.1* transcript analysis of *D. rerio*, nor by *in silico* analysis.

The alternative isoform *OlPax2.1(del35)*, lacks the greater part of a3 helix of PAI domain and ends at a premature stop codon exactly at the beginning of exon 3. Truncated isoforms are not a rare phenomenon, given the fact that approximately 35% of alternatively spliced human transcripts have been found to contain a premature termination codon, rendering them as candidates for non-sense- mediated decay (Green et al., 2003; Lewis et al., 2003). It has been proposed that most low copy number alternative isoforms produced in human cells are likely to be non-functional, therefore we assume that this is also the case for *OlPax2.1(del35)*. Deletion of a3 helix has been observed in one of the Pax6 isoforms in *B. floridae* (Glardon et al., 1998), yet this deletion does not influence downstream translation and hence its functionality. A 32 bp deletion in mouse is responsible for splotch phenotype in mouse (Epstein et al., 1991). In addition, there are accumulating reports about heterozygous deletions of parts of PAI subdomain in general or a3 helix in specific, most of which cause a frame shift and a premature stop codon (Schimmenti et al., 1997; Fletcher et al., 2005) and are correlated with diseases in human (e.g., renal-coloboma syndrome, oligomeganephronia).

In regard to the Pax6 class and the alternative splice isoform *Pax6(5a)*, our analysis showed that an important re-arrangement of coding and non-coding sequences in the region of paired domain took place during evolution. Although conservation of the position of introns has been noted between highly divergent eukaryotes, the number and placement of the majority of introns are dynamically fluctuating during evolution (Hartung et al., 2002; Rogozin et al., 2003). In Hemichordata and Echinodermata, the exon-intron organization does not allow for any type of insertions in the paired domain. In other lineages compared, paired domain is encoded by exons disrupted by one or more introns. Incidents of intron gain and loss as well as intron sliding have been reported for various genes (Hartung et al., 2002), whereas the intron density, i.e., the average number of intron per gene does not necessarily coincide with the position of the genome on the evolutionary tree (Jeffares et al., 2006). We assume that the 5a insertion is characteristic of Gnathostomata.

Introns are required for alternative splicing and alternative splicing increases the size of the proteome, thus increasing the level of complexity in higher eukaryotes. Moreover, introns have been found to harbor many conserved non-coding elements, necessary for gene regulation (Irvine et al., 2008; Bhatia et al., 2014).

Unique isoforms were detected during expression pattern study of *Branchiostoma Pax258* and *Pax46* genes. This is in agreement with *in silico* analysis, during which no splicing events involving the paired domain were predicted. In contrast, new alternative spliced variants were identified for fish species. Previous studies have shown that there is no developmental regulation of paired domain alternative splice forms of *Pax6* and *Pax8*, as opposed to splicing events affecting the C-terminal sequences of Pax8 protein (Kozmik et al., 1993, 1997). In principle, non-constitutive *OlPax2* isoforms are expressed at low levels, therefore at this stage, it is not easy to conclude as to the regulation of these isoforms. Nonetheless there is an indication of a temporal regulation of *OlPax2.1(2a+)* isoform, which requires further investigation.

Acknowledgments

This study was supported by grant LH12047 from the Ministry of Education, Youth and Sports of the Czech Republic and the project "BIOCEV—Biotechnology and Biomedicine Centre of the Academy of Sciences and Charles University" (CZ.1.05/1.1.00/02.0109). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Supplementary Material

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fgene.2015.00228

References

Arseneau, J. R., Laflamme, M., Lewis, S. M., Maïcas, E., and Ouellette, R. J. (2009). Multiple isoforms of *PAX5* are expressed in both lymphomas and normal B-cells. *Br. J. Haematol.* 147, 328–338. doi: 10.1111/j.1365-2141.2009.07859.x

Azuma, N., Tadokoro, K., Asaka, A., Yamada, M., Yamaguchi, Y., Handa, H., et al. (2005). The *Pax6* isoform bearing an alternative spliced exon promotes the development of the neural retinal structure. *Hum. Mol. Genet.* 14, 735–745. doi: 10.1093/hmg/ddi069

Bassham, S., Canestro, C., and Postlethwait, J. (2008). Evolution of developmental roles of *Pax2/5/8* paralogs after independent duplication in urochordate and vertebrate lineages. *BMC Biol.* 6:35. doi: 10.1186/1741-7007-6-35

Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., et al. (2013). GenBank. *Nucleic Acids Res.* 41, D36–D42. doi: 10.1093/nar/gks1195

Bergthorsson, U., Andersson, D. I., and Roth, J. R. (2007). Ohno's dilemma: evolution of new genes under continuous selection. *Proc. Natl. Acad. Sci. U.S.A.* 104, 17004–17009. doi: 10.1073/pnas.0707158104

Bhatia, S., Monahan, J., Ravi, V., Gautier, P., Murdoch, E., Brenner, S., et al. (2014). A survey of ancient conserved non-coding elements in the *PAX6* locus reveals a landscape of interdigitated cis-regulatory archipelagos. *Dev. Biol.* 387, 214–228. doi: 10.1016/j.ydbio.2014.01.007

Blake, J. A., and Ziman, M. R. (2014). *Pax* genes: regulators of lineage specification and progenitor cell maintenance. *Development* 141, 737–751. doi: 10.1242/dev.091785

Bopp, D., Jamet, E., Baumgartner, S., Burri, M., and Noll, M. (1989). Isolation of two tissue-specific *Drosophila* paired box genes, *Pax meso* and *Pax neuro*. *EMBO J.* 8, 3447–3457.

Borson, N. D., Lacy, M. Q., and Wettstein, P. J. (2002). Altered mRNA expression of Pax5 and Blimp-1 in B cells in multiple myeloma. *Blood* 100, 4629–4639. doi: 10.1182/blood.V100.13.4629

Breitling, R., and Gerber, J. K. (2000). Origin of the paired domain. *Dev. Genes Evol.* 210, 644–650. doi: 10.1007/s004270000106

Burge, C., and Karlin, S. (1997). Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* 268, 78–94. doi: 10.1006/jmbi.1997.0951

Busse, A., Rietz, A., Schwartz, S., Thiel, E., and Keilholz, U. (2009). An intron 9 containing splice variant of PAX2. *J. Transl. Med.* 7:36. doi: 10.1186/1479-5876-7-36

Cameron, R. A., Samanta, M., Yuan, A., He, D., and Davidson, E. (2009). SpBase: the sea urchin genome database and web site. *Nucleic Acids Res.* 37, D750–D754. doi: 10.1093/nar/gkn887

Canestro, C., Bassham, S., and Postlethwait, J. (2005). Development of the central nervous system in the larvacean *Oikopleura dioica* and the evolution of the chordate brain. *Dev. Biol.* 285, 298–315. doi: 10.1016/j.ydbio.2005.06.039

Chen, L., Tovar-Corona, J. M., and Urrutia, A. O. (2012). Alternative splicing: a potential source of functional innovation in the eukaryotic genome. *Int. J. Evol. Biol.* 2012:596274. doi: 10.1155/2012/596274

Chi, N., and Epstein, J. A. (2002). Getting your pax straight: pax proteins in development and disease. *Trends Genet.* 18, 41–47. doi: 10.1016/S0168-9525(01)02594-X

Chisholm, A. D., and Horvitz, H. R. (1995). Patterning of the *Caenorhabditis elegans* head region by the *Pax-6* family member *vab-3*. *Nature* 377, 52–55. doi: 10.1038/377052a0

Cinar, H. N., and Chisholm, A. D. (2004). Genetic Analysis of the *Caenorhabditis elegans pax-6* Locus: roles of paired domain-containing and nonpaired domain-containing isoforms. *Genetics* 168, 1307–1322. doi: 10.1534/genetics.104.031724

Cunningham, F., Amode, M. R., Barrell, D., Beal, K., Billis, K., Brent, S., et al. (2015). Ensembl 2015. *Nucleic Acids Res.* 43, D662–D669. doi: 10.1093/nar/gku1010

Czerny, T., Schaffner, G., and Busslinger, M. (1993). DNA sequence recognition by pax proteins: bipartite structure of the paired domain and its binding site. *Genes Dev.* 7, 2048–2061. doi: 10.1101/gad.7.10.2048

De Castro, E., Sigrist, C. J. A., Gattiker, A., Bulliard, V., Langendijk-Genevaux, P. S., Gasteiger, E., et al. (2006). ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res.* 34, W362–W365. doi: 10.1093/nar/gkl124

Dressler, G. R., Deutsch, U., Chowdhury, K., Nornes, H. O., and Gruss, P. (1990). *Pax2*, a new murine paired-box-containing gene and its expression in the developing excretory system. *Development* 109, 787–795.

Eberhard, D., Jimenez, G., Heavey, B., and Busslinger, M. (2000). Transcriptional repression by Pax5 (BSAP) through interaction with corepressors of the Groucho family. *EMBO J.* 19, 2292–2303. doi: 10.1093/emboj/19.10.2292

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Epstein, D. J., Vekemans, M., and Gros, P. (1991). *Splotch* (sp^{2H}), a mutation affecting development of the mouse neural tube, shows a deletion within the paired homeodomain of *Pax-3*. *Cell* 67, 767–774. doi: 10.1016/0092-8674(91)90071-6

Epstein, J. A., Glaser, T., Cai, J., Jepeal, L., Walton, D. S., and Maas, R. L. (1994). Two independent and interactive DNA-binding subdomains of the Pax6 paired

domain are regulated by alternative splicing. *Genes Dev.* 8, 2022–2034. doi: 10.1101/gad.8.17.2022

Escriva, H., Manzon, L., Youson, J., and Laudet, V. (2002). Analysis of lamprey and hagfish genes reveals a complex history of gene duplications during early vertebrate evolution. *Mol. Biol. Evol.* 19, 1440–1450. doi: 10.1093/oxfordjournals.molbev.a004207

Fletcher, J., Hu, M., Berman, Y., Collins, F., Grigg, J., Mciver, M., et al. (2005). Multicystic Dysplastic kidney and variable phenotype in a family with a novel deletion mutation of *PAX2*. *J. Am. Soc. Nephrol.* 16, 2754–2761. doi: 10.1681/ASN.2005030239

Friedrich, M., and Caravas, J. (2011). New insights from hemichordate genomes: prebilaterian origin and parallel modifications in the paired domain of the pax gene eyegone. *J. Exp. Zool. B Mol. Dev. Evol.* 316, 387–392. doi: 10.1002/jez.b.21412

Fu, W., and Noll, M. (1997). The Pax2 homolog sparkling is required for development of cone and pigment cells in the *Drosophila* eye. *Genes Dev.* 11, 2066–2078. doi: 10.1101/gad.11.16.2066

Fuentes, M., Benito, E., Bertrand, S., Paris, M., Mignardot, A., Godoy, L., et al. (2007). Insights into spawning behavior and development of the European amphioxus (*Branchiostoma lanceolatum*). *J. Exp. Zool. B Mol. Dev. Evol.* 308, 484–493. doi: 10.1002/jez.b.21179

Gehring, W. J. (1996). The master control gene for morphogenesis and evolution of the eye. *Genes Cells* 1, 11–15. doi: 10.1046/j.1365-2443.1996.11011.x

Gehring, W. J. (2002). The genetic control of eye development and its implications for the evolution of the various eye-types. *Int. J. Dev. Biol.* 46, 65–73.

Gehring, W. J. (2012). The evolution of vision. *Wiley Interdiscip. Rev. Dev. Biol.* 3, 1–40. doi: 10.1002/wdev.96

Glardon, S., Holland, L. Z., Gehring, W. J., and Holland, N. D. (1998). Isolation and developmental expression of the amphioxus *Pax-6* gene (*AmphiPax-6*): insights into eye and photoreceptor evolution. *Development* 125, 2701–2710.

Glaser, T., Walton, D. S., and Maas, R. L. (1992). Genomic structure, evolutionary conservation and aniridia mutations in the human *PAX6* gene. *Nat. Genet.* 2, 232–239. doi: 10.1038/ng1192-232

Green, R. E., Lewis, B. P., Hillman, R. T., Blanchette, M., Lareau, L. F., Garnett, A. T., et al. (2003). Widespread predicted nonsense-mediated mRNA decay of alternatively-spliced transcripts of human normal and disease genes. *Bioinformatics* 19(Suppl. 1), i118–i121. doi: 10.1093/bioinformatics/btg1015

Grigoriev, I. V., Nordberg, H., Shabalov, I., Aerts, A., Cantor, M., Goodstein, D., et al. (2012). The genome portal of the department of energy joint Genome Institute. *Nucleic Acids Res.* 40, D26–D32. doi: 10.1093/nar/gkr947

Hartung, F., Blattner, F. R., and Puchta, H. (2002). Intron gain and loss in the evolution of the conserved eukaryotic recombination machinery. *Nucleic Acids Res.* 30, 5175–5181. doi: 10.1093/nar/gkf649

Heller, N., and Brandli, A. W. (1997). *Xenopus* Pax-2 displays multiple splice forms during embryogenesis and pronephric kidney development. *Mech. Dev.* 69, 83–104. doi: 10.1016/S0925-4773(97)00158-5

Heller, N., and Brandli, A. W. (1999). *Xenopus Pax-2/5/8* orthologues: novel insights into *Pax* gene evolution and identification of *Pax-8* as the earliest marker for otic and pronephric cell lineages. *Dev. Genet.* 24, 208–219. doi: 10.1002/(sici)1520-6408(1999)24:3/4<208::aid-dvg4>3.0.co;2-j

Hill, A., Boll, W., Ries, C., Warner, L., Osswalt, M., Hill, M., et al. (2010). Origin of Pax and Six gene families in sponges: single *PaxB* and *Six1/2* orthologs in *Chalinula loosanoffi*. *Dev. Biol.* 343, 106–123. doi: 10.1016/j.ydbio.2010.03.010

Holland, L. Z., and Short, S. (2010). Alternative splicing in development and function of chordate endocrine systems: a focus on *Pax* genes. *Integr. Comp. Biol.* 50, 22–34. doi: 10.1093/icb/icq048

Holland, L. Z., and Yu, J. K. (2004). Cephalochordate (amphioxus) embryos: procurement, culture, and basic methods. *Methods Cell. Biol.* 74, 195–215. doi: 10.1016/S0091-679X(04)74009-1

Hoshiyama, D., Iwabe, N., and Miyata, T. (2007). Evolution of the gene families forming the *Pax/Six* regulatory network: isolation of genes from primitive animals and molecular phylogenetic analyses. *FEBS Lett.* 581, 1639–1643. doi: 10.1016/j.febslet.2007.03.027

Howard-Ashby, M., Materna, S. C., Brown, C. T., Chen, L., Cameron, R. A., and Davidson, E. H. (2006). Identification and characterization of homeobox transcription factor genes in *Strongylocentrotus purpuratus*, and their expression in embryonic development. *Dev. Biol.* 300, 74–89. doi: 10.1016/j.ydbio.2006.08.039

Irvine, S. Q., Fonseca, V. C., Zompa, M. A., and Antony, R. (2008). *Cis*-regulatory organization of the *Pax6* gene in the ascidian *Ciona intestinalis*. *Dev. Biol.* 317, 649–659. doi: 10.1016/j.ydbio.2008.01.036

Iwamatsu, T. (2004). Stages of normal development in the medaka *Oryzias latipes*. *Mech. Dev.* 121, 605–618. doi: 10.1016/j.mod.2004.03.012

Jaillon, O., Aury, J. M., Brunet, F., Petit, J. L., Stange-Thomann, N., Mauceli, E., et al. (2004). Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* 431, 946–957. doi: 10.1038/nature03025

Jeffares, D. C., Mourier, T., and Penny, D. (2006). The biology of intron gain and loss. *Trends Genet.* 22, 16–22. doi: 10.1016/j.tig.2005.10.006

Karolchik, D., Barber, G. P., Casper, J., Clawson, H., Cline, M. S., Diekhans, M., et al. (2014). The UCSC Genome Browser database: 2014 update. *Nucleic Acids Res.* 42, D764–D770. doi: 10.1093/nar/gkt1168

Kelemen, O., Convertini, P., Zhang, Z., Wen, Y., Shen, M., Falaleeva, M., et al. (2013). Function of alternative splicing. *Gene* 514, 1–30. doi: 10.1016/j.gene.2012.07.083

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B., and Schilling, T. F. (1995). Stages of embryonic development of the zebrafish. *Dev. Dyn.* 203, 253–310. doi: 10.1002/aja.1002030302

Kleinjan, D. A., Bancewicz, R. M., Gautier, P., Dahm, R., Schonthaler, H. B., Damante, G., et al. (2008). Subfunctionalization of duplicated zebrafish *pax6* genes by *cis*-regulatory divergence. *PLoS Genet.* 4:e29. doi: 10.1371/journal.pgen.0040029

Klimova, L., and Kozmik, Z. (2014). Stage-dependent requirement of neuroretinal Pax6 for lens and retina development. *Development* 141, 1292–1302. doi: 10.1242/dev.098822

Koralewski, T. E., and Krutovsky, K. V. (2011). Evolution of exon-intron structure and alternative splicing. *PLoS ONE* 6:e18055. doi: 10.1371/journal.pone.0018055

Kozmik, Z., Czerny, T., and Busslinger, M. (1997). Alternatively spliced insertions in the paired domain restrict the DNA sequence specificity of Pax6 and Pax8. *EMBO J.* 16, 6793–6803. doi: 10.1093/emboj/16.22.6793

Kozmik, Z., Daube, M., Frei, E., Norman, B., Kos, L., Dishaw, L. J., et al. (2003). Role of *Pax* genes in eye evolution: a cnidarian *PaxB* gene uniting Pax2 and Pax6 functions. *Dev. Cell.* 5, 773–785. doi: 10.1016/S1534-5807(03)00325-3

Kozmik, Z., Holland, N. D., Kalousova, A., Paces, J., Schubert, M., and Holland, L. Z. (1999). Characterization of an amphioxus paired box gene, *AmphiPax2/5/8*: developmental expression patterns in optic support cells, nephridium, thyroid-like structures and pharyngeal gill slits, but not in the midbrain-hindbrain boundary region. *Development* 126, 1295–1304.

Kozmik, Z., Kurzbauer, R., Dorfler, P., and Busslinger, M. (1993). Alternative splicing of *Pax-8* gene transcripts is developmentally regulated and generates isoforms with different transactivation properties. *Mol. Cell. Biol.* 13, 6024–6035. doi: 10.1128/MCB.13.10.6024

Kozmik, Z. (2008). The role of *Pax* genes in eye evolution. *Brain Res. Bull.* 75, 335–339. doi: 10.1016/j.brainresbull.2007.10.046

Krauss, S., Johansen, T., Korzh, V., and Fjose, A. (1991). Expression of the zebrafish paired box gene *pax[zf-b]* during early neurogenesis. *Development* 113, 1193–1206.

Kreslova, J., Holland, L. Z., Schubert, M., Burgtorf, C., Benes, V., and Kozmik, Z. (2002). Functional equivalency of amphioxus and vertebrate Pax258 transcription factors suggests that the activation of mid-hindbrain specific genes in vertebrates occurs via the recruitment of Pax regulatory elements. *Gene* 282, 143–150. doi: 10.1016/S0378-1119(01)00840-X

Kwak, S. J., Vemaraju, S., Moorman, S. J., Zeddies, D., Popper, A. N., and Riley, B. B. (2006). Zebrafish *pax5* regulates development of the utricular macula and vestibular function. *Dev. Dyn.* 235, 3026–3038. doi: 10.1002/dvdy.20961

Lang, D., Powell, S. K., Plummer, R. S., Young, K. P., and Ruggeri, B. A. (2007). *PAX* genes: roles in development, pathophysiology, and cancer. *Biochem. Pharmacol.* 73, 1–14. doi: 10.1016/j.bcp.2006.06.024

Lewis, B. P., Green, R. E., and Brenner, S. E. (2003). Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc. Natl. Acad. Sci. U.S.A* 100, 189–192. doi: 10.1073/pnas.0136770100

Louis, A., Muffato, M., and Roest Crollius, H. (2013). Genomicus: five genome browsers for comparative genomics in eukaryota. *Nucleic Acids Res.* 41, D700–D705. doi: 10.1093/nar/gks1156

Lun, K., and Brand, M. (1998). A series of *no isthmus* (*noi*) alleles of the zebrafish *pax2.1* gene reveals multiple signaling events in development of the midbrain-hindbrain boundary. *Development* 125, 3049–3062.

Maere, S., and Van De Peer, Y. (2010). "Duplicate retention after small- and large-scale duplications," in *Evolution after Gene Duplication*, eds K. Dittmar and D. Liberles (Hoboken, NJ: John Wiley & Sons, Inc.), 31–56.

Nilsen, T. W., and Graveley, B. R. (2010). Expansion of the eukaryotic proteome by alternative splicing. *Nature* 463, 457–463. doi: 10.1038/nature 08909

Noll, M. (1993). Evolution and role of *Pax* genes. *Curr. Opin. Genet. Dev.* 3, 595–605. doi: 10.1016/0959-437X(93)90095-7

Nornes, S., Mikkola, I., Krauss, S., Delghandi, M., Perander, M., and Johansen, T. (1996). Zebrafish *Pax9* encodes two proteins with distinct C-terminal transactivating domains of different potency negatively regulated by adjacent N-terminal sequences. *J. Biol. Chem.* 271, 26914–26923. doi: 10.1074/jbc.271.43.26914

Paixao-Cortes, V. R., Salzano, F. M., and Bortolini, M. C. (2013). Evolutionary history of chordate *PAX* genes: dynamics of change in a complex gene family. *PLoS ONE* 8:e73560. doi: 10.1371/journal.pone.0073560

Pfeffer, P. L., Gerster, T., Lun, K., Brand, M., and Busslinger, M. (1998). Characterization of three novel members of the zebrafish *Pax2/5/8* family: dependency of *Pax5* and *Pax8* expression on the *Pax2.1* (*noi*) function. *Development* 125, 3063–3074.

Poleev, A., Wendler, F., Fickenscher, H., Zannini, M. S., Yaginuma, K., Abbott, C., et al. (1995). Distinct functional properties of three human paired-box-protein, PAX8, isoforms generated by alternative splicing in thyroid, kidney and Wilms' tumors. *Eur. J. Biochem.* 228, 899–911. doi: 10.1111/j.1432-1033.1995.0899m.x

Puschel, A. W., Gruss, P., and Westerfield, M. (1992). Sequence and expression pattern of *pax-6* are highly conserved between zebrafish and mice. *Development* 114, 643–651.

Putnam, N. H., Butts, T., Ferrier, D. E., Furlong, R. F., Hellsten, U., Kawashima, T., et al. (2008). The amphioxus genome and the evolution of the chordate karyotype. *Nature* 453, 1064–1071. doi: 10.1038/nature06967

Ravi, V., Bhatia, S., Gautier, P., Loosli, F., Tay, B. H., Tay, A., et al. (2013). Sequencing of *Pax6* loci from the elephant shark reveals a family of *Pax6* genes in vertebrate genomes, forged by ancient duplications and divergences. *PLoS Genet.* 9:e1003177. doi: 10.1371/journal.pgen.1003177

Reese, M. G., Eeckman, F. H., Kulp, D., and Haussler, D. (1997). Improved splice site detection in Genie. *J. Comput. Biol.* 4, 311–323. doi: 10.1089/cmb.1997.4.311

Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: the European molecular biology open software suite. *Trends Genet.* 16, 276–277. doi: 10.1016/S0168-9525(00)02024-2

Robichaud, G. A., Nardini, M., Laflamme, M., Cuperlovic-Culf, M., and Ouellette, R. J. (2004). Human Pax-5 C-terminal isoforms possess distinct transactivation properties and are differentially modulated in normal and malignant B cells. *J. Biol. Chem.* 279, 49956–49963. doi: 10.1074/jbc.M407171200

Rogozin, I. B., Wolf, Y. I., Sorokin, A. V., Mirkin, B. G., and Koonin, E. V. (2003). Remarkable interkingdom conservation of intron positions and massive, lineage-specific intron loss and gain in eukaryotic evolution. *Curr. Biol.* 13, 1512–1517. doi: 10.1016/S0960-9822(03)00558-X

Schimmenti, L. A., Cunliffe, H. E., Mcnoe, L. A., Ward, T. A., French, M. C., Shim, H. H., et al. (1997). Further delineation of renal-coloboma syndrome in patients with extreme variability of phenotype and identical *PAX2* mutations. *Am. J. Hum. Genet.* 60, 869–878.

Schwartz, S., Zhang, Z., Frazer, K. A., Smit, A., Riemer, C., Bouck, J., et al. (2000). PipMaker–a web server for aligning two genomic DNA sequences. *Genome Res.* 10, 577–586. doi: 10.1101/gr.10.4.577

Short, S., and Holland, L. Z. (2008). The evolution of alternative splicing in the *Pax* family: the view from the Basal chordate amphioxus. *J. Mol. Evol.* 66, 605–620. doi: 10.1007/s00239-008-9113-5

Short, S., Kozmik, Z., and Holland, L. Z. (2012). The function and developmental expression of alternatively spliced isoforms of amphioxus and *Xenopus laevis* *Pax2/5/8* genes: revealing divergence at the invertebrate to vertebrate transition. *J. Exp. Zool. B Mol. Dev. Evol.* 318, 555–571. doi: 10.1002/jez.b.22460

Stephens, R. M., and Schneider, T. D. (1992). Features of spliceosome evolution and function inferred from an analysis of the information at human splice sites. *J. Mol. Biol.* 228, 1124–1136. doi: 10.1016/0022-2836(92)90320-J

Stuart, E. T., Kioussi, C., and Gruss, P. (1994). Mammalian *Pax* genes. *Annu. Rev. Genet.* 28, 219–236. doi: 10.1146/annurev.ge.28.120194.001251

Takatori, N., Butts, T., Candiani, S., Pestarino, M., Ferrier, D. K., Saiga, H., et al. (2008). Comprehensive survey and classification of homeobox genes in the genome of amphioxus, *Branchiostoma floridae*. *Dev. Genes Evol* 218, 579–590. doi: 10.1007/s00427-008-0245-9

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739. doi: 10.1093/molbev/msr121

Tavassoli, K., Ruger, W., and Horst, J. (1997). Alternative splicing in PAX2 generates a new reading frame and an extended conserved coding region at the carboxy terminus. *Hum. Genet.* 101, 371–375. doi: 10.1007/s004390050644

Van De Peer, Y. (2004). *Tetraodon* genome confirms *Takifugu* findings: most fish are ancient polyploids. *Genome Biol.* 5:250. doi: 10.1186/gb-2004-5-12-250

Wada, S., Tokuoka, M., Shoguchi, E., Kobayashi, K., Di Gregorio, A., Spagnuolo, A., et al. (2003). A genomewide survey of developmentally relevant genes in *Ciona intestinalis*. II. Genes for homeobox transcription factors. *Dev. Genes Evol.* 213, 222–234. doi: 10.1007/s00427-003-0321-0

Walther, C., and Gruss, P. (1991). *Pax-6*, a murine paired box gene, is expressed in the developing CNS. *Development* 113, 1435–1449.

Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., et al. (2008a). Alternative isoform regulation in human tissue transcriptomes. *Nature* 456, 470–476. doi: 10.1038/nature07509

Wang, Q., Fang, W. H., Krupinski, J., Kumar, S., Slevin, M., and Kumar, P. (2008b). *Pax* genes in embryogenesis and oncogenesis. *J. Cell. Mol. Med.* 12, 2281–2294. doi: 10.1111/j.1582-4934.2008.00427.x

Ward, T. A., Nebel, A., Reeve, A. E., and Eccles, M. R. (1994). Alternative messenger RNA forms and open reading frames within an additional conserved region of the human *PAX-2* gene. *Cell Growth Differ.* 5, 1015–1021.

Xu, H. E., Rould, M. A., Xu, W., Epstein, J. A., Maas, R. L., and Pabo, C. O. (1999). Crystal structure of the human Pax6 paired domain–DNA complex reveals specific roles for the linker region and carboxy-terminal subdomain in DNA binding. *Genes Dev* 13, 1263–1275. doi: 10.1101/gad.13.10.1263

Xu, W., Rould, M. A., Jun, S., Desplan, C., and Pabo, C. O. (1995). Crystal structure of a paired domain-DNA complex at 2.5 A resolution reveals structural basis for Pax developmental mutations. *Cell* 80, 639–650. doi: 10.1016/0092-8674(95)90518-9

Zwollo, P., Arrieta, H., Ede, K., Molinder, K., Desiderio, S., and Pollock, R. (1997). The *Pax-5* gene is alternatively spliced during B-cell development. *J. Biol. Chem.* 272, 10160–10168. doi: 10.1074/jbc.272.15.10160

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

VII.2.4. **Liegertová M, Pergner J, Kozmiková I, Fabian P, Pombinho AR, Strnad H, Pačes J, Vlček Č, Bartůněk P, Kozmik Z.** Cubozoan genome illuminates functional diversification of opsins and photoreceptor evolution. **Sci Rep. 2015 Jul 8;5:11885 Dev Genes Evol. 2**

www.nature.com/scientificreports

SCIENTIFIC REPORTS

OPEN

Cubozoan genome illuminates functional diversification of opsins and photoreceptor evolution

Received: 10 February 2015
Accepted: 05 June 2015
Published: 08 July 2015

Michaela Liegertová[1,*], Jiří Pergner[1,*], Iryna Kozmiková[1], Peter Fabian[1], Antonio R. Pombinho[2], Hynek Strnad[3], Jan Pačes[3], Čestmír Vlček[3], Petr Bartůněk[2] & Zbyněk Kozmik[1]

Animals sense light primarily by an opsin-based photopigment present in a photoreceptor cell. Cnidaria are arguably the most basal phylum containing a well-developed visual system. The evolutionary history of opsins in the animal kingdom has not yet been resolved. Here, we study the evolution of animal opsins by genome-wide analysis of the cubozoan jellyfish *Tripedalia cystophora*, a cnidarian possessing complex lens-containing eyes and minor photoreceptors. A large number of opsin genes with distinct tissue- and stage-specific expression were identified. Our phylogenetic analysis unequivocally classifies cubozoan opsins as a sister group to c-opsins and documents lineage-specific expansion of the opsin gene repertoire in the cubozoan genome. Functional analyses provided evidence for the use of the Gs-cAMP signaling pathway in a small set of cubozoan opsins, indicating the possibility that the majority of other cubozoan opsins signal via distinct pathways. Additionally, these tests uncovered subtle differences among individual opsins, suggesting possible fine-tuning for specific photoreceptor tasks. Based on phylogenetic, expression and biochemical analysis we propose that rapid lineage- and species-specific duplications of the intron-less opsin genes and their subsequent functional diversification promoted evolution of a large repertoire of both visual and extraocular photoreceptors in cubozoans.

Many animals sense light cues for vision and nonvisual photoreception. Light is captured by an opsin-based photopigment in a photoreceptor cell and leads to a cellular light response through a G protein-mediated phototransduction cascade[1,2]. Opsins are members of the G protein-coupled receptor (GPCR) superfamily; proteins with seven transmembrane helices that are involved in a diverse set of signaling functions. Within the GPCR superfamily, the opsins form a large monophyletic subclass of proteins characterized by a lysine in the seventh transmembrane helix that serves as the attachment site for the chromophore. Functional opsin proteins covalently bind a chromophore, gaining photosensitivity. Opsins are essential molecules in mediating the ability of animals to detect and use light for diverse biological functions and have been discovered in a wide variety of tissue and cell types, signaling through multiple pathways, and carrying out functions beyond image formation[1,3].

Phylogenetic analysis has indicated that four major opsin monophyletic groups can be recognized[1,3–5]. The first group, is comprised of the c-type opsins, the vertebrate visual (transducin-coupled) and non-visual opsin subfamily, the encephalopsins, pinopsins, paprapinopsins, parietopsins and tmt-opsin subfamily and the invertebrate ciliary opsins. The second group, Cnidopsins, is a group consisting of all cnidarian opsins, except for so called Nematostella group 1 and Nematostella group 4 opsins, whose phylogenetic position is still unresolved[4,6]. Cnidopsins are exclusively found among cnidarians and not

[1]Department of Transcriptional Regulation, Institute of Molecular Genetics, Videnska 1083, Prague, CZ-14220, Czech Republic. [2]Department of Cell Differentiation, Institute of Molecular Genetics, Videnska 1083, Prague, CZ-14220, Czech Republic. [3]Department of Genomics and Bioinformatics, Institute of Molecular Genetics, Videnska 1083, Prague, CZ-14220, Czech Republic. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to Z.K. (email: kozmik@img.cas.cz)

present in any other phyla. The third so-called 'r-type' group consists of Gq-coupled invertebrate visual opsins and vertebrate and invertebrate melanopsins and Group 4 opsins contain an assortment of relatively poorly characterized opsin types including, neuropsins, peropsins, and a mixed group of RGRs (retinal G-protein coupled receptors)[7–10]. The distribution of the opsins into these four major groups is supported by analyses of intron arrangements and insertion/deletion events[3] and all groups contain genes found in multiple tissue locations (*e.g.* photoreceptor cells (PRCs) and/or other tissues). Two recently published analyses of opsin phylogeny by Feuda *et al.*[6,11] have shown that in contrast to the findings of a majority of other studies[3,5,10], Cnidarian opsins, including the Nematostella group 1 and Nematostella group 4 opsins, might not be of monophyletic origin, but rather can be divided into three groups, each more closely related to either the c-, r- or Group 4 opsins.

One of the defining characteristics of opsins are the presence of covalently bound chromophores, most commonly 11-cis retinal, that confer light sensitivity to the visual pigment. The chromophore is attached via a Schiff base linkage to a universally conserved lysine in the seventh transmembrane helix. Upon exposure to light, the chromophore undergoes a photoisomerization event to form all-trans retinal, that in turn drives the activation of the photopigment. Aside from this universally conserved lysine, other important amino acid residues can also be found in opsin primary structures. One example is the so-called 'counterion', typically a negatively charged amino acid that is required to interact with and thus raise the pKa of the protonated Schiff base linkage between retinal and lysine, stabilizing the binding of a proton at physiological pH. While vertebrate visual pigments use amino acid-113 as the counterion, position 181 can also be used by a diverse group of opsins containing photoisomerases and Gi/o-coupled pigments, whereas vertebrate melanopsins utilize amino acid position 83[3]. Multiple lines of evidence support the hypothesis that amino acid substitutions in the fourth cytoplasmic loop of duplicated opsins were involved in the origins of novel opsin-G protein interactions[5]. Residues 310-312, encompassing the so-called tripeptide region, were formerly demonstrated by site-directed mutagenesis to mediate opsin-G protein interactions in ciliary opsins[12] and these data were later supported by correlation analyses[5].

Box jellyfish belong to the phylum Cnidaria, arguably the most basal phylum containing a well-developed visual system. It is well known that light affects many behavioral activities of cnidarians, including diel vertical migration, responses to rapid changes in light intensity and reproduction[13]. Their phylogenetic position, simple nervous system and elaborate set of eyes[14] make their visual system of key importance for understanding the early evolution of vision, and also for understanding the biology of box jellyfish[15–18]. The eyes of box jellyfish share many features with those of vertebrates. Morphologically, they are similar by overall design (lens, retina, ciliary photoreceptors)[14,19], and recently, characterization of some molecular components has suggested that the box jellyfish visual system is more closely related to vertebrate than to invertebrate visual systems[20–23]. Photoreceptive organs in Cnidaria have diverse structures, not only between species[24] but within the same species. The cubozoan jellyfish investigated in this study, *Tripedalia cystophora*, has four equally spaced sensory structures called rhopalia, dangling from a stalk and situated within open cavities surrounding the bell. Each rhopalium has six separate eyes. There are two complex, lens-containing eyes (upper lens eye – ULE, and lower lens eye – LLE), one larger than the other, situated at right angles to each other, and two pairs (one pit-shaped, one slit-shaped) of simple ocelli comprising photoreceptors on either side of the complex eyes[25,26]. As the visual fields of individual eyes of the rhopalium partly overlap, *T. cystophora* (as well as other Cubomedusae) has an almost complete view of its surroundings. The lens-containing *T. cystophora* eyes have sophisticated visual optics, similar to molluscs and vertebrates[14,19]. Two opsin genes have so far been identified in cubozoans, one in *T. cystophora*[22], and one in the related species *Carybdea rastonii*[27]. Expression of both these opsins has been detected in eyes of the corresponding species[22,27]. *C. rastonii* opsin was furthermore shown to transfer the light stimulus via the Gs signaling pathway[27].

In the present work, we characterize a complement of 18 opsin genes identified in cubozoan jellyfish *T. cystophora* by the whole-genome analysis. Based on phylogenetic, expression and biochemical analysis we propose that rapid lineage- and species-specific duplications of the intron-less opsin genes and their subsequent functional diversification promoted evolution of both visual and extraocular photoreception in cubozoans.

Results

A large complement of opsin genes are present in *T. cystophora* genome. In addition to previously annotated *T. cystophora* c-opsin[22] we identified 17 other *Tripedalia* opsins (Tcops) sequences. Among those novel sequences, we were able to identify the ortholog (93% sequence identity) of the previously investigated *C. rastonii* opsin (Caryb)[27], designated here in *T. cystophora* as Tcop13 (Fig. S1). Complete coding sequences for all these opsins were obtained by Genome Walking (GenomeWalker, Clontech). All of the eighteen *T. cystophora* opsins are intron-less genes, which show overall sequence homology with other cnidarian opsins as well as to bilaterian rhodopsins. The conserved lysine to which the chromophore 11-cis-retinal binds was found in each of the cloned opsins, suggesting that they are indeed used for photoreception. Next, we investigated the three potential counterion sites at amino-acid position 83, 113, 181 (numbered according to bovine rhodopsin) within the cnidopsins (Fig. S2). Negatively charged amino acids (either glutamic acid/E or aspartic acid/ D) at position 83 was found in more than 50% of the identified cnidopsins, with more than 95% having E/D at position 181. Intriguingly, E/D residues at position 113 were only found in four of the identified *T. cystophora* opsins.

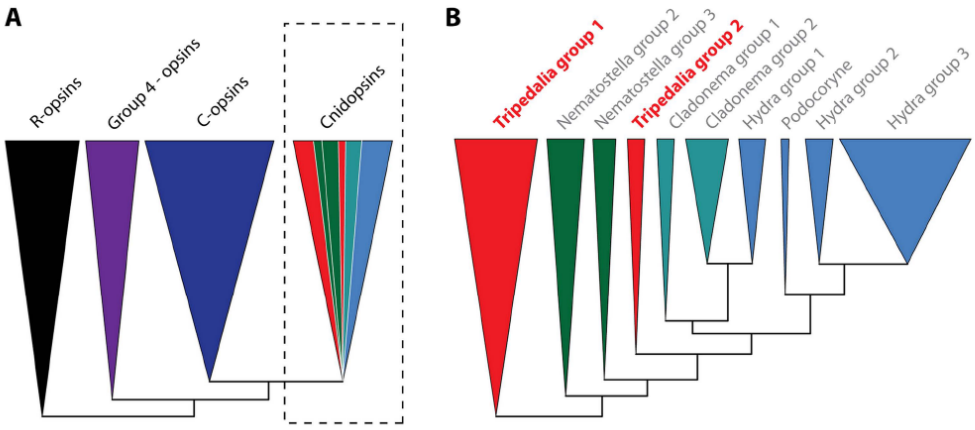


Figure 1. Schematic representation of the opsin phylogenetic analysis of a large set of opsin genes including the cubozoan dataset. A) Phylogenetic analysis performed in this study recovered previously described four major opsin lineages – r-opsins, c-opsins, group 4 opsins and cnidopsins. Herein the C-opsins and cnidopsins form sister groups. (For details see Fig. S3), **B)** Detailed inspection of cnidopsin branch indicates extensive gene duplication and lineage-specific expansion of cnidarian opsins. (For details see Fig. S4).

These were, Tcop8, Tcop12, Tcop15, and Tcop16 (Fig. S2 – red box). However, it is important to note that E/D counterions at position 113 have, to date, not been found in any opsins identified outside of chordates. Next, we investigated the identity of residues 310–312 within all the Tcop sequences. The tripeptides tended to be conserved among closely related cnidopsin groups of each species (Fig. S2) but are apparently not conserved across the cnidarian lineages. In summary, our collective data indicate that a large repertoire of diverse opsins is present in the cubozoan genome, some of which have some intriguing sequence similarities to vertebrate opsins.

Phylogenetic relationships of cubozoan opsins within the opsin gene family. To investigate the relationship between the newly identified Tcops and other known metazoan opsins, we inferred a molecular phylogenetic tree by the maximum likelihood method from a set of 779 opsin protein sequences. Our phylogenetic analysis of this large and diverse set of opsin sequences recovered the four major lineages described in earlier studies[3–5,10], *i.e.* the c-type opsins, the cnidopsins, the r-type opsins, and group 4 opsins. The relationships among the four major lineages in our analyses correlated with those proposed in other recent studies of opsin evolution[4,5,10], however, the statistical support for some of the relationships was weak. Due to such weak branch support, we were unable to exclude the possibility that group 4 and r-type opsins cluster together as sister groups, opposing the c-type opsin and cnidopsin subgroups as has been suggested by Porter *et al.*[3] (based on their phylogeny and the presence of functional unity as bistable pigments of arthropod/cephalopod visual r-type opsins and chicken group 4 neuropsin[28,29]). We found that the relationship between cnidopsins and the c-type opsin subfamily was most strongly supported. All cnidarian opsins except for *Nematostella* group 1 and *Nematostella* group 4 opsins fell within the cnidopsins, group and as was shown by Suga *et al.*[4], the phylogenetic positions of these two groups remained unclear even after more precise Maximum-Likelihood analyses. In the phylogenetic tree of the opsin family, all identified Tcops fell into the cnidopsins subfamily (Fig. 1A, Fig. S3), consistently clustering with the hydrozoan opsins[30]. The cnidopsin group, composed entirely of cnidarian opsins, is the only group lacking representation of broad taxonomic diversity from the major animal phyla. The phylogenic trees presented here (Fig. 1B, Fig. S4) and elsewhere[3,4] are indicative of extensive gene duplications (and diversifications) in each of the cnidarian lineages after their initial split. This latter point is exemplified by the case of Tcops that form two distinct phylogenetic groups: Tc-group 1 and Tc-group 2 (Fig. 1B).

In summary, our phylogenetic analysis unequivocally classifies cubozoan opsins as a sister group to c-type opsins and documents lineage-specific expansion of the opsin gene repertoire in the cubozoan genome.

Functional diversification of opsins in *T. cystophora*: evidence of an apparent dichotomy in G protein-coupled signaling. Sequence analysis and phylogenetic classification provided an important insight into the evolutionary history and possible relationships among opsins. However, these sequence-based approaches do not answer the question whether one or more signaling pathways are being used by *T. cystophora* opsins and do not permit the drawing of conclusions regarding which signaling pathway is coupled to any particular opsin. In order to get a deeper insight into the functional diversification of opsins identified in *T. cystophora* we used a GloSensor™ cAMP HEK293 cell based Gs protein-coupled signaling assay[31] to investigate biochemical properties of all Tripedalia opsins (for description see Material and methods). We used *C. rastonii* opsin (Caryb), shown to activate the Gs-cAMP pathway[27], as a positive control. As heterologous protein expression in cell lines may sometimes prove difficult or even impossible[32], we first checked the expression of the individual opsin genes in GloSensor™ cAMP HEK293 cells by immunofluorescent labeling. The staining revealed that all examined opsins were expressed in GloSensor™ cAMP HEK293 cells and at comparable levels. Moreover the sub-cellular fluorescent signal for opsin was consistently detectable on the cell membranes (Fig. S5 and data not shown), thus confirming successful expression of the Tcop genes. The luciferase activity in GloSensor™ cAMP HEK293 cells, transfected with individual opsin constructs and pre-incubated in the dark with 9-cis retinal, was determined before and after repeated light stimulations (Fig. 2). Light stimulation of cells was in specific cases immediately followed by increased luciferase activity reaching a maximum after several minutes (Tcop5) or 10 minutes (Tcop13, Caryb) and remaining constant for several minutes before decreasing to the basal levels observed prior to illumination (Tcop5) or slightly higher (Caryb; Tcop13). Comparison of the previously characterized Caryb with its ortholog from *T. cystophora*, Tcop13, revealed that both opsins show similar light responses (Fig. 2A). In contrast, medaka (*Oryzias latipes*) opsin RH1, expected to signal via a distinct G protein-coupled pathway (Gi, leading to a cGMP decrease), elicited no increase of luciferase activity in our assay (Fig. 2A), being expressed at comparable levels to those of *T.cystophora* opsins (Fig. S5). Furthermore, no light-dependent stimulation of the Gs protein-coupled assay was detected using the invertebrate r-opsin gene, expected to function via Gq signaling (data not shown). Our assay was, therefore, highly specific for opsins signaling via the Gs-cAMP pathway, but was insensitive to signaling via Gi or Gq. We performed the light response assay several times for the entire set of *T. cystophora* opsins. Of all the opsins examined, only Tcop5 and Tcop13 activated the Gs-cAMP signaling pathway (Fig. 2B–E and data not shown). No convincing light induced opsin-Gs-cAMP response, similar to that of Caryb, Tcop5 and Tcop13, was detected for other Tcops. Tcop5 and Tcop13 sustained enhancement of G protein-coupled pathway signaling after repeated light stimulation. However, we noticed a conspicuous difference in the light response between these two opsins. Tcop5 responded to light faster but with lower intensity (Fig. 2C), whereas the response of Tcop13 was considerably slower, ultimately reaching higher signaling values (Fig. 2D).

To better understand the molecular features of cnidopsins, we focused on the role of the tripeptide in cnidopsin signaling. As stated above, the tripeptide is important for the contact between bovine rhodopsin and its G protein[12]. Accordingly, we replaced the HKQ tripeptide region in Tcop13 with either the tripeptides NKQ, SKS and NRS, originally found in Tcop1 (or bovine rhodopsin), Tcop14 and Tcop18, respectively, none of which activated the Gs signaling cascade in our assay. We expected to observe a loss of Gs cascade activation resulting from the tripeptide mutation. Surprisingly, we found that the tripeptide mutation in Tcop13 did not disrupt Gs activation; rather, it influenced the dynamics of the response to the light stimulation. Specifically, the introduction of the tripeptides NKQ and SKS led to an enhanced and prolonged response, while the introduction of NRS variant caused a massive light response after both single and repeated stimulation (Fig. 2F). Our data show that tripeptide mutation in cnidopsins contributes to subtle tuning of the opsin response to light stimulation, rather than influencing Tcop-G protein coupling *per se*.

In summary, Gs-cAMP signaling characterized only a small set of *T. cystophora* opsins, indicating that the majority of cubozoan opsins likely signal by a distinct and as of yet unidentified signaling pathway. Moreover, the highly sensitive two-dimensional functional assay used here (measuring time response as well as response intensity) uncovered subtle differences among individual opsins, suggesting possible fine tuning for specific photoreceptor tasks.

Phototactic behavior of *T. cystophora* medusa is dependent on Gs signaling. Caryb opsin present in retinas of lensed eyes of *C. rastonii* was previously shown to transfer the light stimulus via the Gs signaling pathway[27]. To investigate whether Tc-group 2 (especially Tcop13), that signal via Gs (see above), serve as important visual pigments in *T. cystophora*, we performed a phototaxis behavioral assay in the absence or presence of the pharmacological compound NF449. NF449 was originally identified as a selective suppressor of the Gs signaling pathway, with limited effect on the prototypical Gi/Go- and Gq-coupled receptors pathways[33]. Positive phototaxis in *T. cystophora* medusa was significantly decreased after treatment with NF449. Although, a variable response was detected in response to white light depending on NF449 concentration and timing of the treatments (Fig. 3 + video), most probably due to reversible inhibition of the Gs signaling pathway by NF449. The number of treated animals exhibiting a phototactic response 5 minutes after the treatment was: 5 ± 5% in samples treated with 100 μM NF449 and 0% in samples treated with 1mM. The number of responding medusae treated with 100 μM NF449 after 3 h rose to 95 ± 5% and the number of medusae treated with 1 mM NF449 rose to 10%.

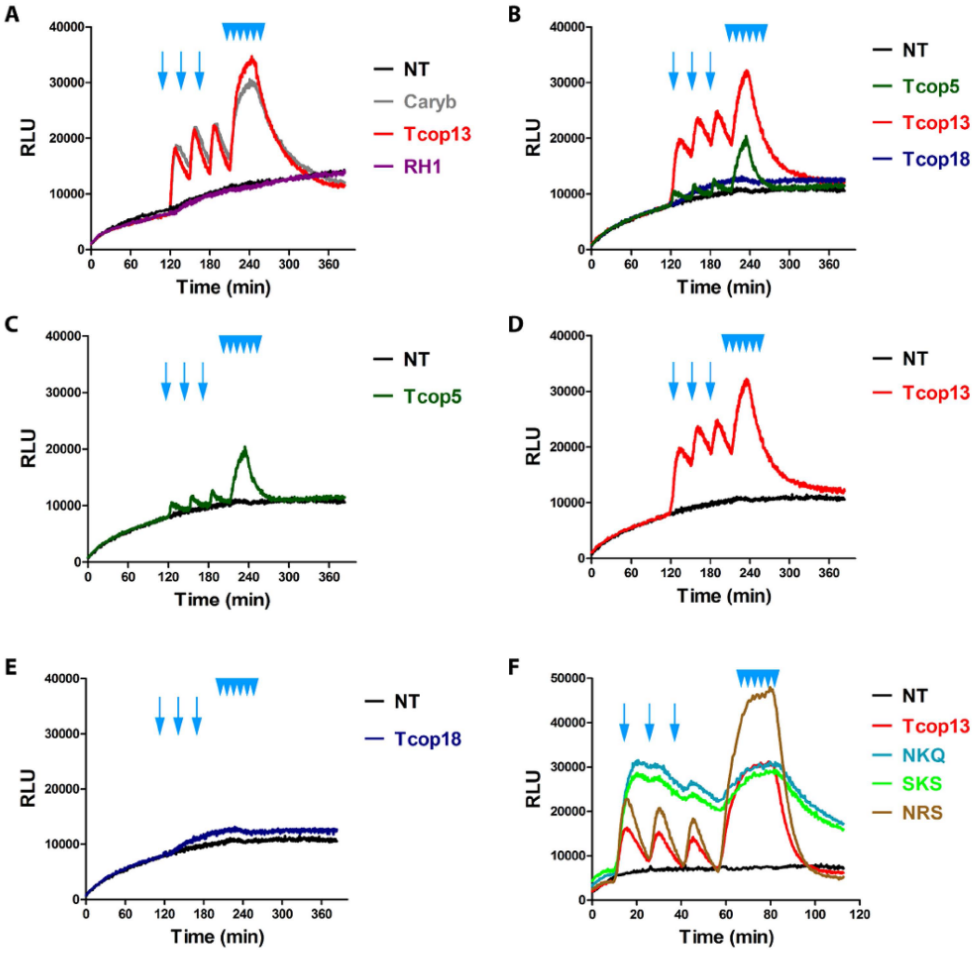


Figure 2. Opsin-Gs-cAMP assay. Light activation of opsin-Gs-cAMP pathway by selected opsins. GloSensor™-20F cAMP HEK293 cells (Promega) were transfected with expression vectors encoding genes for different opsins, treated and stimulated with light, as described in Materials and Methods. Arrows represent simple light pulses, multiple arrowheads represent repeated stimulation. Each graph represents a mean of triplicates for every sample. **A)** Previously reported Gs-cAMP pathway stimulating opsin from *C. rastonii* (Caryb)[27] showed ability to increase the cAMP level in our setup (visualized with cAMP-dependent luciferase activity). The exact homolog of Caryb from *T. cystophora* Tcop13 showed a highly similar response in our assay. Opsin RH1 from medaka, expected to signal via Gt leading to cGMP decrease, showed no change in luciferase activity. **B)–E)** Examples of different Tcop light responses. Tcop5 showed faster and weaker activation of the Gs-cAMP pathway than Tcop13. Tcop18 did not activate the Gs-cAMP pathway. **F)** Analysis of tripeptide activity in Tcop13 was performed. Tcop13 tripeptide HKQ was replaced with tripeptides NKQ, SKS and NRS (originally found in opsins Tcop1 or bovine rhodopsin, Tcop14 and Tcop18 – none of which activated the Gs cascade). Tripeptide mutation did not disrupt Gs activation by Tcop13, but influenced length or sensitivity of Tcop13 response to light stimulation. NT – non-transfected cells used as negative control; Caryb – signal for cells transfected with a vector expressing opsin from *C. rastonii*, used as positive control; RH1 – signal for cells transfected with a vector expressing opsin RH1 from medaka fish *Oryzias latipes*, used as negative control; Tcop5, Tcop13, Tcop18 – signal for cells transfected with vectors expressing opsins from *T. cystophora* - Tcop5, Tcop13 or Tcop18, respectively; NKQ, SKS, NRS – Tcop13 original tripeptide HKQ replaced with tripeptides NKQ, SKS or NRS.

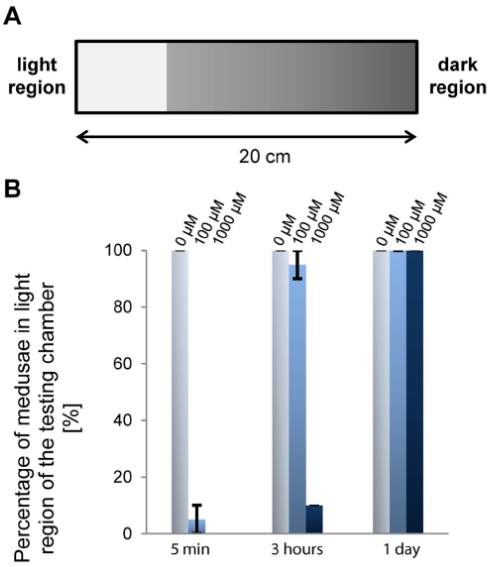


Figure 3. Test of *T. cystophora* medusa phototaxis after NF449 treatment. **A**) Schematic representation of the testing chamber. **B**) Statistical analysis of the light response of *T. cystophora* after NF449 (Gαs inhibitor) treatments (0 µM, 100 µM, 1 mM). Bars represent the percentage of phototactic medusae in given time point.

However, after 24 hours the decrease in photosensory response was no longer present. We observed 100% phototactic response in untreated animals (0 µM NF449) after 5 min, 3 h, 24 h intervals (Fig. 3B). Thus, the pharmacological inhibition of Tc-group 2 opsins abrogates positive phototactic movement of cubozoan medusa.

Opsin gene expression analysis reveals both redundancy and specialization. The large complement of opsins found in the *T. cystophora* genome raises the possibility of their differential tissue-specific or stage-specific utilization. To investigate the expression patterns of *T. cystophora* opsins, we first analyzed mRNA isolated from different jellyfish life stages and dissected adult tissues by real-time qRT-PCR analysis. The normalized expression levels, relative to Rpl32 levels, of specific opsin genes in each dissected adult body part was calculated relative to that observed in the rhopalium (set at 1.0) and plotted (Fig. 4). Relevant opsin expression data were also represented as a heat map showing z-score of Tcops expression in different *T. cystophora* body parts. (Fig. S6A). All Tcop genes were found to be expressed at the mRNA level in the rhopalium. Moreover, for the majority of opsins (Tcop1, Tcop3–7, Tcop9, Tcop11, Tcop13, Tcop16, Tcop18), the rhopalium was the tissue exhibiting the highest detected expression. Other opsins were most highly expressed in the male gonads (Tcop2, Tcop10, Tcop17), tentacles (Tcop12, Tcop14) or manumbrium (Tcop8, Tcop15). For a better gene-to-gene comparison within the rhopalium, the results were plotted separately (Fig. S7). The expression data in the adult tissues identified common/over-lapping sites of expression most probably reflecting a common gene origin. Nevertheless, a clear tendency for specialization was apparent as very large differences in relative expression levels and/or unique sites of expression were detected.

Next, we investigated opsin gene expression during the life cycle of *T. cystophora*. To this end, mRNAs from non-pigmented larva, pigmented larva (larval eye-containing stage), vegetatively grown polyp, four stages of a metamorphosing polyp (stages 3 and 4 containing a developing rhopalia) and medusae were isolated and subjected to qRT-PCR. The expression levels of all individual opsins for each *T. cystophora* life stage relative to the juvenile medusa expression (set as 1.0) were then calculated (Fig. 5 and Fig. S6B). The results revealed two consistent features. Firstly, opsins whose expression was detected in the adult rhopalium, sharply increased their expression during the polyp metamorphosis, coincident with the emergence of the developing rhopalia structure. Secondly, many Tc-group 1 opsins were highly expressed in the pigmented (eye-containing) larval stage, contrasting with the expression of Tc-group 2 opsins (established as Gs-coupled receptors with a major functional role in the adult lens-containing eye, see above), that were notably absent at this stage.

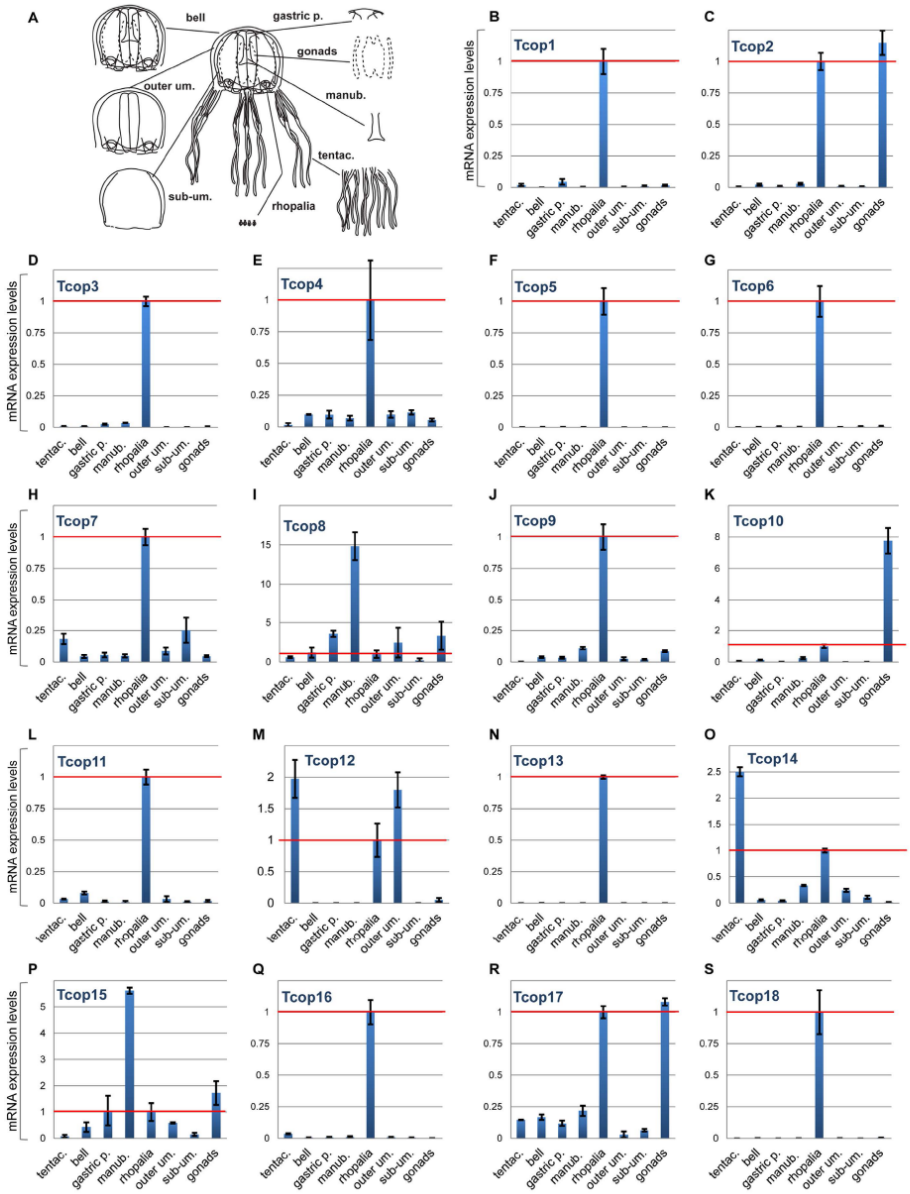


Figure 4. mRNA expression levels of *T. cystophora* opsins in dissected body parts of adult jellyfish. A) For real-time PCR analysis, medusae were dissected into eight body parts: tentacles (tentac.), manubrium (manb.), male gonads, gastric pouch (gastric p.), bell, outer umbrella (outer um.), sub-umbrella (sub-um.) and rhopalia. **B–S)** mRNA expression level of opsins for each dissected body part relative to the rhopalium expression (1.0 – red line). y – axis: relative mRNA expression level, x – axis: analyzed body parts.

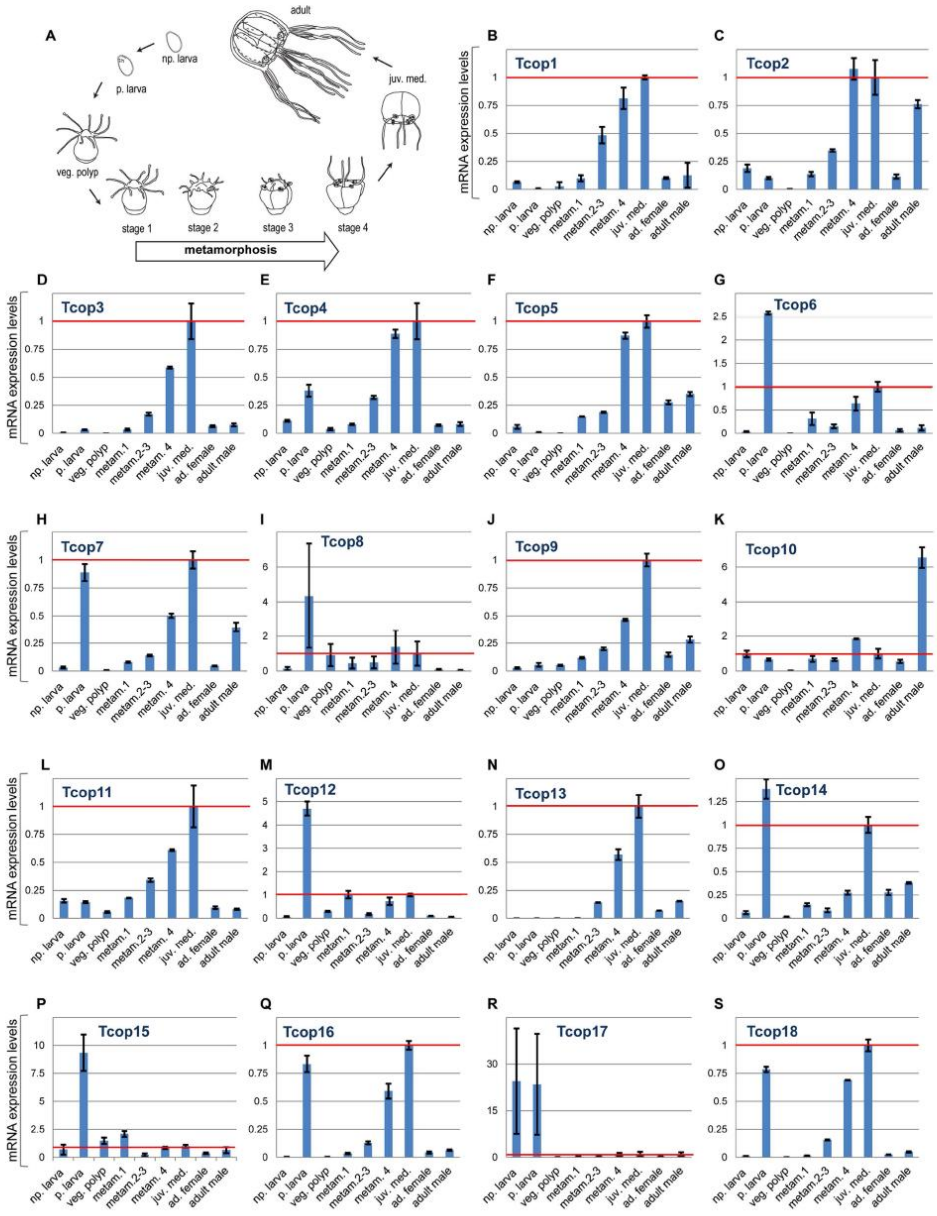


Figure 5. mRNA expression levels of *T. cystophora* opsins in different life stages . A) For real-time PCR analysis, animals of nine subsequent life stages were collected: non-pigmented larva (np. larva), pigmented larva (p. larva), vegetative polyp (veg.polyp), three polyp-to-medusa metamorphosing stages (metam1, 2–3, 4), juvenile medusa (juv. med.), adult female (ad. female) and adult male. B-S) mRNA expression levels of opsins for each life stage relative to the juvenile medusa expression (1.0 – red line). x – axis: analyzed stages, y – axis: relative mRNA expression level.

To gain further insight into the possibly diverse roles of Tc-group 1 and Tc-group 2 opsins in the various cubozoan eyes (Fig. 6A–C), we also analyzed the expression of key representatives of each opsin type by immunohistochemical staining (IHC) *in situ*. Accordingly, we generated a specific antibody against Tcop13 and performed co-staining with another antibody raised against Tcop18[22] on cryo-sectioned rhopalia. We found that both Tcop13 and Tcop18 were found to be co-expressed in the retinas of *T. cystophora* ULE and LLE in distinct patterns (Fig. 6D–P). We also discovered that *T. cystophora* retinas contain at least two distinct photoreceptor types: ciliary photoreceptor type-A that express Tcop18 not restricted only to the cilia but rather expressed more broadly within the whole cell body (Fig. S8) plus ciliary photoreceptor type-B (expressing Tcop13 in the receptor cell cilia). Both opsins were also distinctly expressed in the developing lens eyes of the newly metamorphosed *T. cystophora* medusae; however, only Tcop18 was detected in the developing eyes of another Carybdeid jellyfish, *Alatina marsupialis* (Fig. S9). Another difference in the apparent utilization of Tc-group 1 and Tc-group 2 opsins was revealed by further analysis of their expression in the two lesser eye types, slit and pit eyes, whose morphology has been thoroughly studied[26]. Only Tcop18 (but not Tcop13) was found to be expressed in the pit and slit ocelli of *T. cystophora*; both types of ocelli thus seem to be formed exclusively from type-A photoreceptors based on the opsin type expression (Fig. S10). However, some molecular features are shared by the PRCs of lesser eye types with those of complex lensed eyes. For example, all PRCs in *T. cystophora* contain two different screening pigments, dark pigment and a white pigment, first described in *Chiropsella bronzie*[24], that becomes conspicuous in polarization microscopy (Fig. S11). It is important to stress that the Tc-group 1 opsin Tcop18 is the only known opsin to be expressed in the lesser eyes thus far. All phylogenetic, biochemical and gene expression data are schematically summarized in Fig. 7.

Discussion

A Scenario for intron-less retrogene-derived cnidarian opsin expansion. The current results highlight distinct features of intron-less genes in vertebrates. It appears that many intron-less genes are evolutionary innovations, so their formation, at least in part, via reverse transcription-mediated mechanisms, could be an important route of evolution of tissue-specific functions in animals[34,35].

The lack of introns is typical for most members of the giant GPCR gene family and it has been proposed that many G-protein-coupled receptors are derived and amplified from a single intron-less common progenitor that was encoded by a retrogene (a DNA gene copied into genome by reverse transcription of an RNA transcript)[36,37]. Interestingly the vertebrate rhodopsin GPCRs, that are widespread phylogenetically and abundantly expressed, contain four introns in highly conserved positions[38]. However, most of the cnidarian opsins thus far annotated are intron-less genes, although at least one opsin in anthozoan *Acropora digitifera*, CNOP2, has been characterized with two introns[39]. Astonishingly, the first of these intron matches, in position and phase, with the first intron of bovine rhodopsin (Fig. S12). Such examples of the first introns to be located in cnidarian opsins are moderately short and have conventional GT-AG donor and acceptor splice sites and thus it appears that this intron was already present in an opsin gene present in the last common ancestor of eumetazoa. Accordingly, intron-less opsin genes appear to be a Cnidarian feature, with the original variant of the gene most probably being lost in medusozoans. Similarly, intron-less opsin genes were previously identified in two cephalopod species[40] and in genome of teleost fish[41], in both cases probably derived from introns containing opsin genes by retrotransposition. Based on these facts and our data, we propose the scenario (Fig. S13) that cnidarian intron-less opsins are retrogenes derived from an ancient eumetazoan ciliary-like opsin with introns. This hypothesis is supported by the phylogenetic relationship of c-like opsins and cnidopsins (Fig. 1A) and by the fact that both variants of the opsin gene (with or without introns) are still present in basal anthozoa (Fig. S12). Once an intron-less opsin gene was present in cnidarian genome, subsequent rapid lineage- and species-specific duplications resulted in a variety of opsins. This process provided the substrate for the evolution of cnidarian photoreception, be it either extraocularly or in sophisticated cubozoan eyes.

Gene duplications and their subsequent divergence play an important part in the evolution of novel gene functions[42]. Our data show that in *T. cystophora* genome, each of the opsins has been duplicated at least once, and several have undergone multiple rounds of duplication (Fig. S4). Theory suggests that duplicated genes can be lost rapidly[43], but the spectrally diverse aquatic environment (such as the margins of mangrove lagoons naturally inhabited by *T. cystophora*) could provide strong selective pressures on the opsin genes, and thus, photoreception evolution. Photoreception tuning through opsin sequence evolution might therefore be a result of sensory adaptation to this rich environment of spectral light.

In both medaka and zebrafish the opsin gene diversity in the genome is high, similarly as in the genome of *T. cystophora*. Subtype opsin genes in medaka and zebrafish are closely linked and are clearly products of local gene duplications[44]. Tandem duplication appears to be the most common mode of opsin gene family expansion in fishes[45]. Gene duplication followed by amino acid substitutions at key tuning sites played an important role in generating a diverse set of fish opsin genes. It is probable that similar mechanisms of opsin gene repertoire expansion occurred in the case of cnidaria (evolutionary convergence), where the opsin genes, being relatively short and intron-less, were even more rapidly duplicated and subsequently functionally diversified (see Fig. 8 for schematical representation).

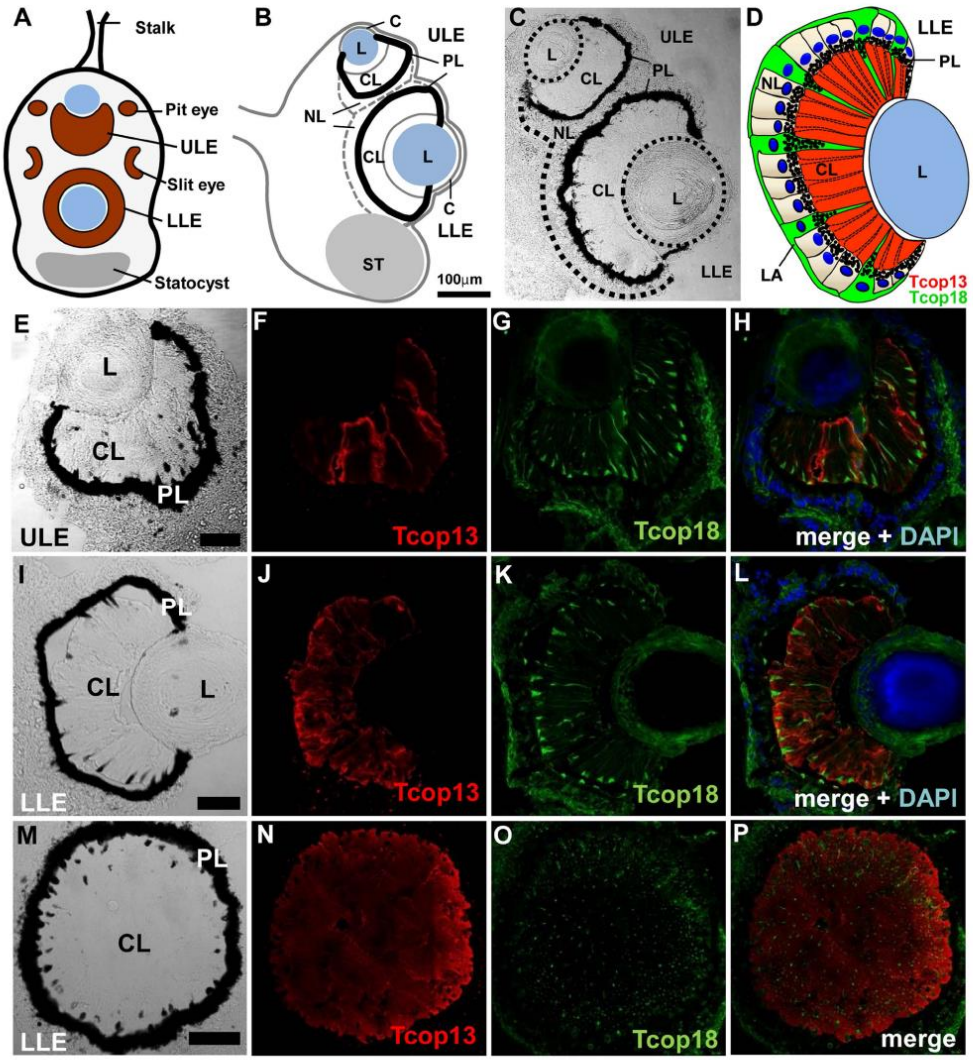


Figure 6. Visual organs of *T. cystophora* and immunohistochemical localization of Tcop13 and Tcop18. A) Schematic diagram of the rhopalium. The large (LLE) and small (ULE) complex eyes lie along the medial line, while the pit and slit ocelli are paired laterally. B) Schematic diagram of rhopalium sagittal plane (adapted from O´Connor 2009). C) Sagittal section through the rhopalium. Upper (ULE) and lower (LLE) lens eyes contain the typical components of camera-type eyes: a cornea (C), a lens (L), and a retina consisting of a ciliary layer (CL), a pigment layer (PL) and a neural layer (NL). St – statocyst, S – stalk. D) Schematic representation of the lens eye retina. The ciliary layer (CL) is dominated by the ciliary segments of type-B receptor cells (red). Scattered among the type-B receptor cells are the cone-shaped projections of type-A photoreceptor cells (green). In the neural layer (NL), both receptors types have their cell bodies with nuclei (dark blue); only type-A receptor cell bodies are positive for opsin signal. Projections of type-A photoreceptor cell bodies create a compact layer (LA) surrounding the whole retina. E–H) Confocal images of immuno-histochemical staining for Tcop13 (red), Tcop18 (green), DAPI (blue) in the upper lens eye (ULE). I-L) Large camera-type eye (LLE) retina longitudinal section. M–P) Large camera-type eye retina transverse section. (Scale bars: 50µm).

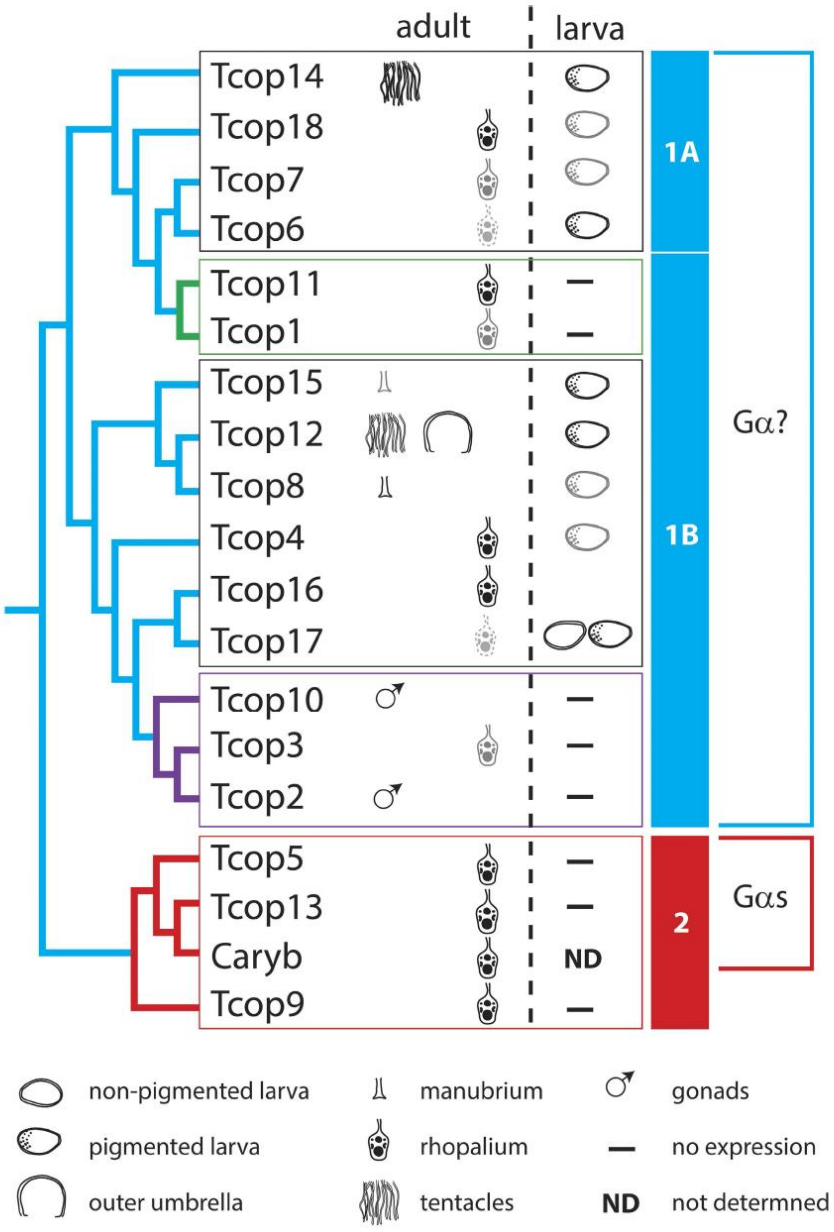


Figure 7. Schematic representation of opsin expression patterns according to their phylogenetic relationship. *T. cystophora* opsins can be classified into two groups, a probable more ancient Tc-group 1 opsin, with a broader expression pattern, and Tc-group 2 – rhopalium-specific opsins. The size and shade intensity of the symbols corresponds with the level of expression. Green coloured box and branches represent rhopalia specific Tc group 1A opsins. Purple coloured box and branches represent male specific Tc group 1B opsins. Red coloured box and branches represent rhopalia specific Tc-group 2 opsins.

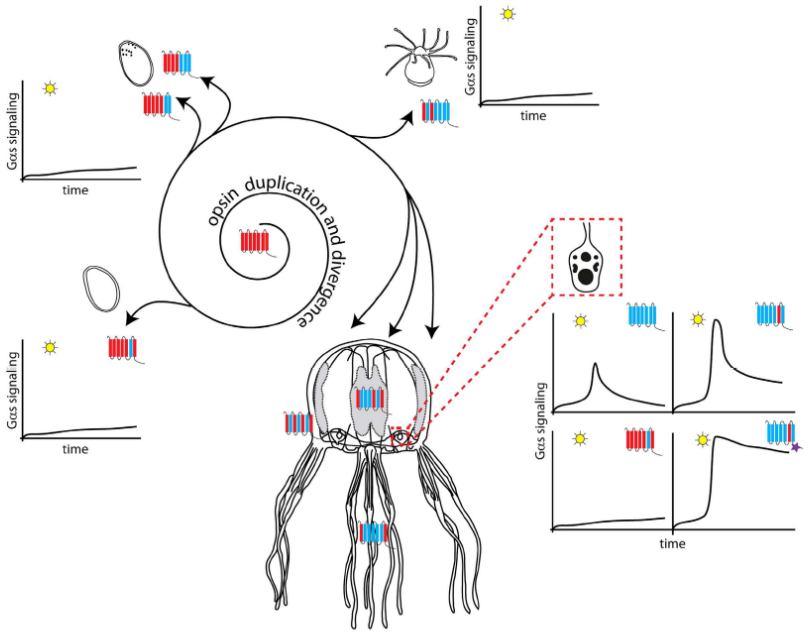


Figure 8. Possible scenario for expansion and functional diversification of opsins in *T. cystophora*. Our data and data from other studies[39,60] show that Cnidarian intron-less opsins might have been derived from an ancient eumetazoan ciliary-like opsin containing introns by retro-transposition. Once anchored in the genome the ancient cnidopsin gene underwent several rounds of duplication, diversification and sensitivity tuning. Individual opsins were thus accommodated for distinct functions in diverse tissue photoreceptors - ocular, extraocular and larval. These opsins differ in stage- or tissue-expression, primary structure and also in subsequent cellular signaling – either via Gs-cAMP pathway or other G-protein pathways. For further information see Discussion.

Rhopalia-specific opsin expression in *T. cystophora*. Cubozoa have relatively simple nervous systems consisting of a nerve net and a ring nerve. The latter has extensions forming ganglia and connections with the radial nerves and rhopalia. Morphological and electrophysiological studies have shown that a significant part of the CNS of cubomedusae is situated within the rhopalia[25,46,47]. In addition to numerous ciliated photoreceptors within the retinas of all six eyes, each rhopalium houses over 1,000 neurons of which approximately 500 are retina-associated. Each rhopalium also contains a group of pacemaker neurons that regulate swimming movements through the direct control of neuronal activity in the motor nerve net, and thus individual rhopalium facilitates various behaviors such as obstacle avoidance or light-shaft attraction enabling them to remain in close proximity to prey gathered in beams of light passing through open parts of the mangrove canopy. This behavioral regulation is most probably influenced by the visual input received by each rhopalium[48].

Based on our mRNA (Fig. 4) and protein (Fig. 6) expression profiles, many of the opsin genes identified here are expressed in rhopalia. Since it is not easy to determine the physiological relevance of a given gene just based on the level of its mRNA expression, we appreciated the finding that Tcop18, with 100 times lower level of mRNA transcripts compared to Tcop13 (Fig. S7), is significantly expressed on protein level (Fig. 6). Based on this fact, we suggest that all Tc-group 2 opsins, plus at least one opsin from each subclass of Tc-group 1 is rhopalium-specific. Moreover, the real-time PCR analysis has revealed that all of the rhopalium-expressed opsins are dramatically up-regulated when the rhopalia are formed during the polyp-to-medusa metamorphosis (Fig. 5). We thus propose that the *Tripedalia* rhopalium is a complex organ integrating and processing multiple light cues, gained though a diverse set of opsins, and transforming these signals into various behavioral responses.

Retina-specific opsin expression in *T. cystophora*. In addition to the extensive real-time PCR expression analysis, we paid special attention to the IHC analysis of retina specific opsins expression in *Tripedlia* rhopalia. The cubozoan lens-containing eyes have a thin cornea (made of monociliated

epithelial cells), a spherical cellular lens, a thin vitreous space, and a hemisphere-shaped everse retina with pigmented photoreceptors of the ciliary type, as judged from their ultra-structural morphology[24,49,50].

A previous study identified three types of photoreceptors in the lensed eyes of *T. cystophora* on the basis of differences in the morphology of their sensory cilium and microvillar organization[50]. In contrast, other studies[51–53] supported the interpretation that there was only a single basic morphological type of photoreceptor in cubozoan lensed eyes. Our IHC data support the first interpretation, showing that there are at least two types of PRC (each with markedly different opsin expression profiles) in the lensed eyes of *T. cystophora*. Both cell types have three distinct segments, giving rise to three retinal layers: 1. a thick layer of receptor-cell cilia formed from type-B PRCs (expressing Tcop13) and cone-shaped projections from type-A PRCs (expressing Tcop18), creating the ciliary layer; 2. a thin pigment layer where both receptor cells types are densely pigmented; 3. a neural layer containing nucleated cell bodies of both types of receptor cells.

The ciliary layer is dominated by the ciliary segments of type-B receptor cells. The cilia extend from the pigment layer to the vitreous space. From the ciliary membrane, microvilli extend, partly as bundles of parallel microvilli and partly as a disorganized tangle (as shown in another cubozoan jellyfish *Ch. bronzie*[24]). The microvilli make up the majority of the volume of the ciliary layer. Scattered among the type-B receptor cells are the cone-shaped projections of type-A photoreceptor cells partially filled with screening pigment granules. These cones run parallel to the ciliary trunks of the type-B sensory cells. In the neural layer, the type-A receptor cells have their cell bodies with nuclei, and they are also positive for Tcop18 protein expression (Fig. 6E–P). Projections of type-A photoreceptor cell bodies create a compact layer surrounding the whole retina.

It has been previously suggested that the lens eye photoreceptors utilize a different photopigment from those of the pit eyes and slit eyes[52]. According to dominant mRNA and protein levels and strict retinal specificity, we consider Tcop13 the main visual opsin of *T. cystophora* complex lens eyes. On the other hand, Tcop18 (also expressed in lens eyes) appears to be the main visual opsin in the lesser eyes. Our IHC data show, that retinas of both eye types (lens and lesser eyes) express different opsin combinations (various combinations of Tcop13 and Tcop18) according to their task (another level of visual tuning). The expression of rhopalium-specific opsins surely does not only involve photoreceptors of the retina, as some of the retina-associated neurons will most probably prove to be photosensitive as well, given our qRT-PCR analysis. This possibility should be resolved in the future by detailed IHC analysis assaying other Tcops expression.

Tissue-specific and larval opsins. Eyes are not the only means of photoreception in the Cnidarians, as many species lack distinct ocular structures yet exhibit specific photic behaviors. In these animals, photosensitivity is mediated through extraocular PRCs. Extraocular photosensitivity, is widespread throughout the animal kingdom, in both invertebrates and vertebrates[54,55]. The extraocular photosensitive cells are not organized into a complex organ such as ocelli or lens eyes. Instead, these cells are solitary or grouped and are scattered or localized throughout the animal body. Identification of the cells involved in extraocular photodetection has often proved difficult, but in some animals, neurons, epithelial cells, and muscle cells have been shown to be photosensitive[54,56–58]. Intriguingly, an ancient opsin-mediated phototransduction pathway and a previously unknown layer of sensory complexity in the control of cnidocyte discharge in cnidarian *Hydra magnipapillata* was reported very recently[59]. These various extraocular photoreceptors function as light detectors, informing the animal of the presence of light, measuring light intensity, and activating rhythmic behaviors as well as other physiological processes.

Our extensive qRT-PCR analysis (Figs 4, 5, S6, S7) (with support from the phylogenetic data) of different developmental stages and tissues revealed that the *T. cystophora* opsins can be classified into two groups, the probably more ancient Tc-group 1 opsins and Tc-group 2 rhopalium-specific opsins (Fig. 7). Tc-group 1 opsins tend to have broader expression. The broadest tissue- and stage-specific expression distribution is visible in Tc-group 1B, with Tcop2 and Tcop10 being male gonad-specific and other Tcops expressed in tissues such as bell, tentacles or manubrium. Both sub-groups 1A and 1B show a trend for increasing tissue/organ specificity of opsins after subsequent duplications.

More than a half of the opsins from those two subgroups were detected (at least in small amounts) in planula larvae, with Tcop17 (and probably Tcop6) being larva-specific. Such a variety of larval opsins is astonishing considering that to date only three larval opsins have been reported from reef corals[60]. Planula larvae have an extremely simple organization with no nervous system at all. Their only advanced feature is the presence of 10–15 pigment-cup ocelli, evenly spaced across the posterior half of the larval ectoderm. The ocelli are single-cell structures containing a cup of screening pigment filled with presumably photosensory microvilli. These morphologically rhabdomeric-like photoreceptors have no neural connections to any other cells, but each has a well-developed motor-cilium, appearing to be the only means by which light can control the behavior of the larva[61].

Our analysis implies that Cnidarians extensively utilize opsins not only for visual but also for extraocular photosensitivity. Revisiting the possible diversity of Tcops tissue/stage specific expression by IHC protein expression analysis and physiological studies could shed more light on their use for various behavioral tasks.

Phototransduction by cubozoan opsins. To investigate the coupling partner of *T. cystophora* opsins we performed an opsin-Gs-cAMP coupling assay. Our data revealed that the Gs-cAMP pathway[27] is used by opsin genes from Tc-group 2. Moreover our behavioral test showed for the first time that the opsin-Gs-cAMP cascade is functionally connected with vision guided behaviour. However, we were unable to obtain any light-mediated activation of signal transduction via this pathway for Tc-group 1 opsins. We propose that opsins that did not signal in our assay either use different G-protein pathways, as recently proposed in reef corals[60], act as photoisomerases or for unknown reason do not signal in our cell-line based assay, but nonetheless use Gs signaling cascade under natural conditions. The later possibility is, however, in our opinion very unlikely, because we saw comparable expression of all Tcops on the cell membranes of our test cell system and moreover not even a repeated flash stimulation lead to any response. However, we do acknowledge the possibility, that some of examined Tcops did not fold properly in mammalian cells used in the assay and thus were unable to signal. In some cases we did record slight increases in luciferase signal (like the one in Fig. 2E for Tcop18), however this phenomenon also appeared in some wells containing cells transfected with control opsins (not signaling though Gs-cAMP cascade) or non-transfected cells. This slight increase in luciferase activity was probably connected with non-opsin-specific changes in cellular metabolism during experiment (note that increase in luciferase activity in Fig. 2E starts 140 minutes from the beginning of the experiment and does not seem to be connected with light stimulation). Clearly the future identification of the actual Gα subunit coupled to Tc-group 1 opsins is going to be necessary to understanding if *T. cystophora* possess at least two independent photosystems, thus providing another level for the functional divergence of the identified opsins. Another interesting feature of our assay is the time-course of response of Tcop5, Tcop13 and Caryb transfected cells to light stimulus, reaching the peak in the order of minutes. This phenomenon is probably not caused by the slow light response of the opsins themselves, but rather indicative of the slow kinetics of the recombinant cAMP-sensitive luciferase expression in GloSensor™ cAMP HEK293 cells. In a study by Koyanagi *et al.*[62], the use of a similar assay led to peak of response in order of minutes even in the case of bovine rhodopsin, which is known to respond to light stimuli in other direct assay systems within millisecond time periods[63].

Future structure-function studies of prototypical cubozoan group 2 opsin is highly warranted. It would be interesting to find out whether any of the proposed E/D counterions are indeed used by *T. cystophora* opsins. Likewise, the significance of various tripeptide variants found among *T. cystophora* opsins awaits further experimental interrogation. Our data so far point to variable sensitivity and bleaching properties of individual opsins depending on their primary amino acid sequence. Based on their expression and conserved amino acid sequence at key positions, we assume that all Tcops described here are functional opsins, but as mentioned earlier, this remains to be confirmed by other analysis (IHC expression, identification of Tc-group 1 signalling cascade).

In summary, our data suggest that the expansion and diversification of the opsin gene family in cubozoans has allowed fine tuning and optimal photopigment function.

In summary, a detailed expression analysis uncovered both redundancy and specialization in the utilization of the opsin gene repertoire. On the one hand, multiple opsins with presumably similar molecular characteristics are apparently utilized in the same stage/tissue. On the other hand, a clear tendency to establish unique expression patterns exists both within the opsin subfamilies (Tc-group 1 and Tc-group 2) and between the two subfamilies. Remarkably, retina photoreceptors of lens-containing eyes express opsins most probably utilizing at least two distinct signaling pathways.

Materials and Methods

Jellyfish collection and culture. Adult *T. cystophora* were collected from the mangroves of La Parguerra, Puerto Rico. Laboratory cultures were established using settling larvae and artificial seawater. Settled larvae metamorphosed into young polyps. Young polyps were transformed into budding (asexually reproducing) polyps by feeding with Artemia once a week. Polyps were stimulated into metamorphosis (transformation into free swimming medusa) by incubation at 28°C. Polyps and young medusa were both maintained at 26 °C. All stages were collected for (opsin) expression pattern analysis (RT-PCR) and juvenile medusa also for rhopalium IHC.

Isolation of *Tripedalia cystophora* opsin genes. *Tripedalia cystophora* genomic DNA shotgun sequencing was performed on the GS FLX Titanium platform (454 Life Sciences, Roche). Pyrosequencing resulted in 1,952,068 reads (about 7×10^8 bases) with average read length of 360 bp. Assembly generated 134,683 of all contigs containing 790,111 (40.5%) reads. Assembly was done by program Newbler, version 2.3 (Roche). Resulting contigs were combined with singleton reads to produce a complete contig database. The database was subjected to similarity search by the FASTA[64] program using a wide range of homologous opsin proteins from other cnidarian and bilaterian species. FASTA search provided hits corresponding to short stretches of assumed Tripedalia opsin protein sequences. Full opsin genes sequences were obtained by using the Genome Walking strategy (Genome Walker, Clontech). Opsin sequences were deposited in GeneBank (accession numbers: JQ968416 -JQ968432).
(Primers in Supplement -T1)

Molecular phylogeny. To investigate the relationship between the cnidarian opsins and bilaterian opsins, we inferred a molecular phylogenetic tree by the maximum likelihood (ML) method implemented in PhyML 3.0[65] with LG substitution model[66]. Support for internal nodes was assessed using Approximate Likelihood-Ratio Test for Branches[67].

Dataset. Opsin protein sequences were acquired as described by Porter *et al.*[3]; however, incomplete sequences were discarded from the analysis and other 26 *Nematostella vectensis* annotated opsins were added to the dataset. In order to root the phylogenetic tree, 22 non-opsin GPCRs from the human genome were used as outgroups. The resulting dataset of 801 (779 opsin plus 22 non-opsin) transcripts plus genome trace opsin sequences was aligned using ClustalX[68] under default parameters and trimmed by eye in BioEdit. For phylogenetic analyses, only the 7-transmembrane region including intervening inter- and extra-cellular domains was included, as it was difficult to ascertain homology of N- and C- termini due to sequence length variation and lack of conservation across genes. The molecular phylogenetic tree of the opsin family was inferred from an alignment of 226 amino acids long (after N- and C-termini exclusion) opsin sequences. (Sequences in Supplement -T2)

Quantitative RT-PCR. RNA from indicated stages or dissected adult *T. cystophora* tissues was isolated using TRIZOL reagent (Invitrogen). Contaminating genomic DNA was removed by DNAse digestion and RNA repurification on RNeasy Micro columns (Qiagen) according to the manufacturer's protocol. The same amounts of RNA from each sample were used for reverse transcription using VILO cDNA kit (Invitrogen). Primers for qPCR were designed using Primer 3 software (see Supplementary Table 1 for sequences of primers). The qPCR was performed in LightCycler 2.0 System using LightCycler® 480 DNA SYBR Green I Master kit (Roche Diagnostics, Germany) according to the standard manufacturer's protocol. Target genes (Tcop1-Tcop18) and the housekeeping gene (Rpl32) were measured under the same conditions from the same cDNA. Results were analyzed by LightCycler software and crossing point values (Cp) were further determined as an average of Cp values from all replicates and normalized by Cp values of the housekeeping gene (so called deltaCp values). The results show relative normalized gene expression. Statistical significance of changes in the mRNA level of target genes between different samples were calculated by a Student´s t-test. For other data reproduction, heat map from z-scores (Standard scores) of deltaCp values for target genes (Tcop1-Tcop18) expression in different *T. cystophora* tissues was constructed. Z-score representation was obtained in R statistical environment with Bioconductor package.

Generation and verification of antibodies. An antibody directed against Tcop13 c-opsin was prepared by immunization of mice as follows. The C-terminal region of *c-opsin* corresponding to amino acids 281-330 (*NPIIYCFLHKQFRRAVLRGVCGRIVGGNAIAPSSTGVEPGQTLGGGAAES*; primers in Supplement -T1) was cloned into the expression vector pET42, expressed in BL21(DE3)RIPL cells (Stratagene), and purified by *Ni*-NTA Agarose Beads (QIAGEN). Purified protein was used as antigen for mouse immunization. Human kidney HEK293 cells were transfected with EGFP_C1-c-opsin (amino acids 281–330) expression vector by using FuGENE®6 reagent (Roche). Total extracts were prepared from c-opsin-transfected cells and mock-transfected cells and were analyzed by Western blotting by using anti-c-opsin mouse serum and chemiluminescent detection kit (Pierce).

Tissue collection and histology. Jellyfish were fixed in 4% paraformaldehyde (PFA), cryoprotected in 30% sucrose overnight at 4°C, and embedded and frozen in OCT (Tissue Freezing Medium, Jung). Horizontal frozen sections were prepared with a 8–12 μm thickness. The cryosections were washed three times in PBS and subsequently immuno-stained with an antibody.

Immunohistochemistry. The cryosections were refixed in 4% PFA for 10 min, washed three times with PBS, permeabilized with PBT (PBS + 0.1% Tween 20) for 15 min, and blocked in 10% BSA in PBT for 30 min. The primary antibodies were diluted in 1% BSA in PBT (1:500), incubated overnight at 4 °C, washed three times with PBS, and incubated with secondary antibodies in 1% BSA in PBT (1:500). The sections were counterstained with DAPI and mounted. Primary antibodies used were: anti-Tcop18[22], anti-Tcop13, and anti-acetylated tubulin (Sigma). The following secondary antibodies were used: Alexa Fluor 488- or 594-conjugated goat anti-mouse or anti-rabbit IgG (Molecular Probes).

Construction of opsin-expressing vectors. The expression vector pcDNA3.1 + 1D4 for opsin gene production in mammalian cells was prepared as follows. The sequence for BamHI restriction site followed by the sequence of 1D4 epitope tag from bovine rhodopsin was introduced into multiple cloning site of pcDNA 3.1+ vector (Clontech) through KpnI and EcoRI sites. Opsin cDNA of box jellyfish *C. rastonii* (GeneBank AB435549), kindly provided Dr. Koyanagi, was amplified from the vector by PCR and cloned into pcDNA3.1 + 1D4 vector using BamHI and HindIII cloning sites. The opsins of box jellyfish *T. cystophora*, which are all intron-less, were amplified by PCR from genomic DNA and cloned

into pcDNA 3.1 + 1D4 vector either via BamHI and HindIII or BamHI and KpnI cloning sites. All the constructs were verified by standard sequencing techniques before use.

Immunofluorescent staining of GloSensor™ cAMP HEK293 cells. GloSensor™ cAMP HEK293 cells (Promega) ($2.5*10^3$) were seeded onto coverslips and transfected with FuGene HD (ROCHE). The next day, cells were washed with PBS and fixed with 4% paraformaldehyde (PFA) for 10 minutes. Fixed cells were permeabilized with 0.1% Triton X-100 for 10 minutes and blocked with 10% BSA in 1x PBS with 0.1% Tween 20 for 1 hour. A mouse monoclonal antibody raised against 1D4 epitope (Millipore Chemicon MAB5356), at a concentration of 1:250, was used in conjunction with a secondary antibody conjugated with Alexa Fluor 488 to immuno-stain expressed opsins. Cells were mounted in Mowiol®. Fluorescent images were captured using a Leica SP5 confocal microscope.

Light response assays. GloSensor™ cAMP HEK293 cells (Promega) ($10*10^3$) were plated into a solid white 96-well plate in L15 CO_2-independent medium with phenol red (Gibco) and 10% serum and incubated overnight at 37 °C, 0.3% CO_2. The cells were transfected the next day with plasmids expressing opsin genes using FuGENE® HD Transfection Reagent (ROCHE). Immuno-fluorescent staining revealed a transfection efficiency of 50% using this method. All procedures following transfection of the cells with the various opsin receptors were carried out in dim red light. Six hours post transfection 9-cis retinal (Sigma-Aldrich) was added to a final concentration of 10 mM. The cells were then kept overnight in an incubator (37 °C; 0.3% CO_2). Next day the cells were removed from the incubator and left to equilibrate for 30 minutes at room temperature. Beetle luciferin potassium salt (Synchem) reconstituted in 10 mM HEPES buffer was added to the cells to a final concentration of 3 mM. The cells were then placed in a top-read Envision plate reader with ultra-sensitive luminescence model. Luciferase activity was measured for 2 hours with 0.1 second resolution and cycles of every 1 minute to determine the luciferin uptake. Cells were then subjected to three pulses of light stimulation using repeated flashes from a Nikon speed-light SB-600 electronic camera flash (5 flashes, 1 flash/ second in each pulse, ~40000 lumen/m2 per flash) followed by recovery periods of 30 minutes when Raw Luminescence Units (RLU) were recorded. After the third measurement, the cells were stimulated with seven light pulses with periods of 3 minutes (5 flashes, 1 flash/ second in each pulse). Luminescence was recorded between pulses (0.1 second resolution, 15 seconds per cycle) and another 120 minutes after the last pulse (0.1 seconds resolution, 30 seconds per cycle). The experiment for the tripeptide mutation was performed in a similar way with minor changes. The entire experiment was performed at 37 °C, which led to faster response of cells to the light stimulation. Three pulses (5 flashes, 1 flash/ second in each pulse) followed by recovery of 15 minutes were applied. The following repeated stimulation was done with 30 light pulses (1 flash) with periods of 30 seconds. Luminescence was measured another 30 minutes after the last pulse. Luminescence recordings were analyzed with Microsoft Office Excel. All experiments comprised cells plated and treated in triplicate. Prism (Graphpad) software was used for all statistical analyses.

***T. cystophora* phototaxis test.** All behavioral tests were performed at room temperature (22 °C). Phototaxis experiments were performed in an aquarium-like testing chamber (20 × 5 × 5 cm) with one illuminated side (Fig. 7a). To test the effect of suramin analog 4,4',4'',4'''-(carbonylbis(imino-5,1,3-benzenetriylbis (carbonylimino)))tetrakis-benzene-1,3-disulfonic acid (NF449 - Calbiochem) on the *T. cystophora* phototactic behavior, we incubated 3-day-old medusae in 1 ml of artificial seawater with NF449 at a final concentration of either 0 μM, 100 μM, and 1 mM for 30 minutes under artificial day light. Medusae were then washed with artificial seawater, placed into the dark part of the testing chamber and tested for phototactic behavior. The number of medusae that reached the light region after 5 minutes, 3 hours and 24 hours was counted and compared to the number of animals from the untreated control group.

References

1. Terakita, A. The opsins. *Genome Biol* **6**, 213, doi: 10.1186/gb-2005-6-3-213 (2005).
2. Sakmar, T. P., Menon, S. T., Marin, E. P. & Awad, E. S. Rhodopsin: insights from recent structural studies. *Annual Review of Biophysics and Biomolecular Structure* **31**, 443-484, doi: 10.1146/annurev.biophys.31.082901.134348 (2002).
3. Porter, M. L. *et al.* Shedding new light on opsin evolution. *Proc Biol Sci* **279**, 3–14, doi: 10.1098/rspb.2011.1819 (2012).
4. Suga, H., Schmid, V. & Gehring, W. J. Evolution and functional diversity of jellyfish opsins. *Current Biology* **18**, 51–55, doi: 10.1016/j.cub.2007.11.059 (2008).
5. Plachetzki, D. C., Degnan, B. M. & Oakley, T. H. The origins of novel protein interactions during animal opsin evolution. *PLoS One* **2**, e1054, doi: 10.1371/journal.pone.0001054 (2007).
6. Feuda, R., Hamilton, S. C., McInerney, J. O. & Pisani, D. Metazoan opsin evolution reveals a simple route to animal vision. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 18868–18872, doi: 10.1073/pnas.1204609109 (2012).
7. Kojima, D. *et al.* A novel Go-mediated phototransduction cascade in scallop visual cells. *Journal of Biological Chemistry* **272**, 22979–22982 (1997).
8. Koyanagi, M., Terakita, A., Kubokawa, K. & Shichida, Y. Amphioxus homologs of Go-coupled rhodopsin and peropsin having 11-cis- and all-trans-retinals as their chromophores. *FEBS Letters* **531**, 525–528 (2002).
9. Hara, T. & Hara, R. Rhodopsin and retinochrome in the squid retina. *Nature* **214**, 573–575 (1967).
10. Plachetzki, D. C., Fong, C. R. & Oakley, T. H. The evolution of phototransduction from an ancestral cyclic nucleotide gated pathway. *Proc Biol Sci* **277**, 1963–1969, doi: 10.1098/rspb.2009.1797 (2010).
11. Feuda, R., R.-S. O., Oakley, T. H. & Pisani, D. The comb jelly opsins and the origins of animal phototransduction. *Genome Biology and Evolution* **6**, 1964-*1971*, doi: 10.1093/gbe/evu154 (2014).

12. Marin, E. P. *et al.* The amino terminus of the fourth cytoplasmic loop of rhodopsin modulates rhodopsin-transducin interaction. *Journal of Biological Chemistry* **275**, 1930–1936 (2000).
13. Martin, V. J. Photoreceptors of cnidarians. *Can. J. Zool.* **80**, 1703–1722, doi: 10.1139/Z02-136 (2002).
14. Nilsson, D. E., Gislen, L., Coates, M. M., Skogh, C. & Garm, A. Advanced optics in a jellyfish eye. *Nature* **435**, 201–205, doi: 10.1038/nature03484 (2005).
15. Coates, M. M. Visual ecology and functional morphology of cubozoa (cnidaria). *Integr Comp Biol* **43**, 542–548, doi: 10.1093/icb/43.4.542 (2003).
16. O'Connor, M., Nilsson, D. E. & Garm, A. Temporal properties of the lens eyes of the box jellyfish Tripedalia cystophora. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* **196**, 213–220, doi: 10.1007/s00359-010-0506-8 (2010).
17. Garm, A., Oskarsson, M. & Nilsson, D. E. Box jellyfish use terrestrial visual cues for navigation. *Current Biology* **21**, 798–803, doi: 10.1016/j.cub.2011.03.054 (2011).
18. Petie, R., Garm, A. & Nilsson, D. E. Visual control of steering in the box jellyfish Tripedalia cystophora. *Journal of Experimental Biology* **214**, 2809–2815, doi: 10.1242/jeb.057190 (2011).
19. Land, M. F. & Nilsson, D. E. Animal eyes. *Oxford, Oxford University Press* (2002).
20. Kozmik, Z. *et al.* Role of Pax genes in eye evolution: a cnidarian PaxB gene uniting Pax2 and Pax6 functions. *Dev Cell* **5**, 773–785 (2003).
21. Piatigorsky, J. & Kozmik, Z. Cubozoan jellyfish: an Evo/Devo model for eyes and other sensory systems. *International Journal of Developmental Biology* **48**, 719-729, doi: 10.1387/ijdb.041851jp (2004).
22. Kozmik, Z. *et al.* Assembly of the cnidarian camera-type eye from vertebrate-like components. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 8989–8993, doi: 10.1073/pnas.0800388105 (2008).
23. Kozmik, Z. *et al.* Cubozoan crystallins: evidence for convergent evolution of pax regulatory sequences. *Evol Dev* **10**, 52–61, doi: 10.1111/j.1525-142X.2007.00213.x (2008).
24. O'Connor, M., Garm, A. & Nilsson, D. E. Structure and optics of the eyes of the box jellyfish Chiropsella bronzie. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* **195**, 557–569, doi: 10.1007/s00359-009-0431-x (2009).
25. Parkefelt, L., Skogh, C., Nilsson, D. E. & Ekstrom, P. Bilateral symmetric organization of neural elements in the visual system of a coelenterate, Tripedalia cystophora (Cubozoa). *Journal of Comparative Neurology* **492**, 251–262, doi: 10.1002/cne.20658 (2005).
26. Garm, A., Andersson, F. & Nilsson, D. E. Unique structure and optics of the lesser eyes of the box jellyfish Tripedalia cystophora. *Vision Research* **48**, 1061–1073, doi: 10.1016/j.visres.2008.01.019 (2008).
27. Koyanagi, M. *et al.* Jellyfish vision starts with cAMP signaling mediated by opsin-G(s) cascade. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 15576–15580, doi: 10.1073/pnas.0806215105 (2008).
28. Panda, S. *et al.* Illumination of the melanopsin signaling pathway. *Science* **307**, 600–604, doi: 10.1126/science.1105121 (2005).
29. Yamashita, T. *et al.* Opn5 is a UV-sensitive bistable pigment that couples with Gi subtype of G protein. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 22084–22089, doi: 10.1073/pnas.1012498107 (2010).
30. Bridge, D., Cunningham, C. W., Schierwater, B., DeSalle, R. & Buss, L. W. Class-level relationships in the phylum Cnidaria: evidence from mitochondrial genome structure. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 8750–8753 (1992).
31. Bailes, H. J., Z. L. & Lucas, R. J. Reproducible and sustained regulation of Gαs signalling using a metazoan opsin as an optogenetic tool. *PLOS one* **7**, doi: 10.1371/journal.pone.0030774 (2012).
32. Koyanagi, M., Kubokawa, K., Tsukamoto, H., Shichida, Y. & Terakita, A. Cephalochordate melanopsin: evolutionary linkage between invertebrate visual cells and vertebrate photosensitive retinal ganglion cells. *Current Biology* **15**, 1065–1069, doi: 10.1016/j.cub.2005.04.063 (2005).
33. Hohenegger, M. *et al.* Gsalpha-selective G protein antagonists. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 346–351 (1998).
34. Shabalina, S. A. *et al.* Distinct patterns of expression and evolution of intronless and intron-containing mammalian genes. *Molecular biology and evolution* **27**, 1745–1749, doi: 10.1093/molbev/msq086 (2010).
35. Zou, M., Guo, B. & He, S. The roles and evolutionary patterns of intronless genes in deuterostomes. *Comparative and functional genomics* **2011**, 680673, doi: 10.1155/2011/680673 (2011).
36. Gentles, A. J. & Karlin, S. Why are human G-protein-coupled receptors predominantly intronless? *Trends in genetics : TIG* **15**, 47–49 (1999).
37. Brosius, J. Many G-protein-coupled receptors are encoded by retrogenes. *Trends in genetics : TIG* **15**, 304–305 (1999).
38. Fridmanis, D., Fredriksson, R., Kapa, I., Schioth, H. B. & Klovins, J. Formation of new genes explains lower intron density in mammalian Rhodopsin G protein-coupled receptors. *Molecular phylogenetics and evolution* **43**, 864–880, doi: 10.1016/j.ympev.2006.11.007 (2007).
39. UCSC, G. B. G. *Opsin evolution: update blog http://genomewiki.ucsc.edu, (Date of acces: 20/05/2012).*
40. Morris, A. B. J. & Hunt, D. M. The molecular basis of a spectral shift in the rhodopsin of two species of squid from different photic enviroments. *Proceedings of the Royal Society of London B* **254**, 233–240 (1993).
41. Fitzgibbon, J., H. A., Slobodvanyuk, S. J., Bellingham, J., Bowmaker, J. K. & Hunt, D. M. The rhodopsin-encoding gene of bony fish lacks introns. *Gene* **164**, 273–277 (1995).
42. Innan, H. & Kondrashov, F. The evolution of gene duplications: classifying and distinguishing between models. *Nature reviews. Genetics* **11**, 97–108, doi: 10.1038/nrg2689 (2010).
43. Lynch, M. & Conery, J. S. The evolutionary fate and consequences of duplicate genes. *Science (New York, N.Y.)* **290**, 1151–1155 (2000).
44. Matsumoto, Y., Fukamachi, S., Mitani, H. & Kawamura, S. Functional characterization of visual opsin repertoire in Medaka (Oryzias latipes). *Gene* **371**, 268–278, doi: 10.1016/j.gene.2005.12.005 (2006).
45. Rennison, D. J., Owens, G. L. & Taylor, J. S. Opsin gene duplication and divergence in ray-finned fish. *Molecular phylogenetics and evolution* **62**, 986–1008, doi: 10.1016/j.ympev.2011.11.030 (2012).
46. Skogh, C., Garm, A., Nilsson, D. E. & Ekstrom, P. Bilaterally symmetrical rhopalial nervous system of the box jellyfish Tripedalia cystophora. *Journal of Morphology* **267**, 1391–1405, doi: 10.1002/jmor.10472 (2006).
47. Garm, A., Ekstrom, P., Boudes, M. & Nilsson, D. E. Rhopalia are integrated parts of the central nervous system in box jellyfish. *Cell Tissue Res* **325**, 333–343, doi: 10.1007/s00441-005-0134-8 (2006).
48. Stockl, A. L., Petie, R. & Nilsson, D. E. Setting the pace: new insights into central pattern generator interactions in box jellyfish swimming. *PLoS One* **6**, e27201, doi: 10.1371/journal.pone.0027201 (2011).
49. Yamasu, T. & Yoshida, M. Fine structure of complex ocelli of a cubomedusan, Tamoya bursaria Haeckel. *Cell and Tissue Research* **170**, 325–339 (1976).
50. Laska, G., Hundgen, M. Morphologie und ultrastruktur der lichtsinnesorgane von Tripedalia cystophora Conant (Cnidaria, Cubozoa). *Zool Jb Anat* **108**, 107–123 (1982).
51. O'Connor, M. *et al.* Visual pigment in the lens eyes of the box jellyfish Chiropsella bronzie. *Proc Biol Sci* **277**, 1843–1848, doi: 10.1098/rspb.2009.2248 (2010).

52. Ekstrom, P., Garm, A., Palsson, J., Vihtelic, T. S. & Nilsson, D. E. Immunohistochemical evidence for multiple photosystems in box jellyfish. *Cell and Tissue Research* **333,** 115–124, doi: 10.1007/s00441-008-0614-8 (2008).
53. Garm, A. & Ekstrom, P. Evidence for multiple photosystems in jellyfish. *Int Rev Cell Mol Biol* **280,** 41–78, doi: 10.1016/s1937-6448(10)80002-4 (2010).
54. Yoshida, M. Extraocular photoreception. *Handbook of sensory physiology* **VII/6A**, 582–640 (1979).
55. Taddei-Ferretti, C. & Musio, C. Photobehaviour of Hydra (Cnidaria, Hydrozoa) and correlated mechanisms: a case of extraocular photosensitivity. *Journal of Photochemistry and Photobiology. B, Biology* **55,** 88–101 (2000).
56. Arkett, S. A. & Spencer, A. N. Neuronal mechanism of a hydromedusan shadow reflex. I. Identified reflex components and sequence of events *J. Comp. Physiol. A* **159,** 201–213 (1986).
57. Sawyer, S. J., D. H. B. & Shick, M. Neurophysiological correlates of the behavioral response to light in the sea anemone Anthopleura elegantissima. *Biol. Bull. (Woods Hole, Mass.)* **186,** 195–201 (1994).
58. Musio, C. *et al.* First identification and localization of a visual pigment in Hydra (Cnidaria, Hydrozoa). *J Comp Physiol A* **187,** 79–81 (2001).
59. Plachetzki, D. C., Fong, C. R. & Oakley, T. H. Cnidocyte discharge is regulated by light and opsin-mediated phototransduction. *BMC Biol* **10**, 17, doi: 10.1186/1741-7007-10-17 (2012).
60. Mason, B. *et al.* Evidence for multiple phototransduction pathways in a reef-building coral. *PLoS One* **7**, e50371, doi: 10.1371/journal.pone.0050371 (2012).
61. Nordstrom, K., Wallen, R., Seymour, J. & Nilsson, D. A simple visual system without neurons in jellyfish larvae. *Proc Biol Sci* **270,** 2349–2354, doi: 10.1098/rspb.2003.2504 (2003).
62. Koyanagi, M.T. E., Nagata, T., Tsukamoto, H. & Terakita, A. Homologs of vertebrate Opn3 potentially serve as a light sensor in nonphotoreceptive tissue. *Proceeding of the National Academy of Sciences of the United States of America* **110,** 4998–5003, doi: 10.1073/pnas.1219416110 (2013).
63. Heck, M. S. S., Maretzki, D., Bartl, F. J., Ritter, E., Palczewski, K. & Hofmann, K. P. Signaling States of Rhodopsin: FORMATION OF THE STORAGE FORM, METARHODOPSIN III, FROM ACTIVE METARHODOPSIN II. *Journal of Biological Chemistry* **278,** 3162–3169 (2003).
64. Pearson, W. R. & Lipman, D. J. Improved tools for biological sequence comparison. *Proceedings of the National Academy of Sciences of the United States of America* **85,** 2444–2448 (1988).
65. Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* **59,** 307–321, doi: 10.1093/sysbio/syq010 (2010).
66. Le, S. Q. & Gascuel, O. An improved general amino acid replacement matrix. *Molecular Biology and Evolution* **25,** 1307–1320, doi: 10.1093/molbev/msn067 (2008).
67. Anisimova, M. & Gascuel, O. Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. *Syst Biol* **55,** 539–552, doi: 10.1080/10635150600755453 (2006).
68. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23,** 2947–2948, doi: 10.1093/bioinformatics/btm404 (2007).

Acknowledgements

We are grateful to Dr. Koyanagi for providing cDNA encoding *C. rastonii* opsin and Sarka Takacova and Dr. Alex Bruce for manuscript proofreading. This study was supported by the Grant Agency of the Czech Republic (P305/10/2141) to Z.K. by the Ministry of Education, Youth and Sports of the Czech Republic (LM2011022) to P.B. and by IMG institutional support RVO68378050.

Author Contributions

Conceived and designed the experiments: M.L., J.P., I.K., P.F., P.B., C.V. and Z.K. Performed the experiments: M.L., J.P., I.K., P.F., A.P. and Z.K. Analyzed the data: M.L., J.P., I.K., P.F., A.P., H.S., J.P.a. and Z.K. Wrote the paper: M.L., J.P. and Z.K.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Liegertová, M. *et al.* Cubozoan genome illuminates functional diversification of opsins and photoreceptor evolution. *Sci. Rep.* **5**, 11885; doi: 10.1038/srep11885 (2015).

www.nature.com/scientificreports

OPEN

Corrigendum: Cubozoan genome illuminates functional diversification of opsins and photoreceptor evolution

Michaela Liegertová, Jiří Pergner, Iryna Kozmiková, Peter Fabian, Antonio R. Pombinho, Hynek Strnad, Jan Pačes, Čestmír Vlček, Petr Bartůněk & Zbyněk Kozmik

Scientific Reports 5:11885; doi: 10.1038/srep11885; published online 08 July 2015; updated on 24 September 2015

This Article contains a typographical error in the grant number in the Acknowledgements section.

"This study was supported by the Grant Agency of the Czech Republic (P305/10/2141) to Z.K. by the Ministry of Education, Youth and Sports of the Czech Republic (LM2011022) to P.B. and by IMG institutional support RVO68378050".

should read:

"This study was supported by the Grant Agency of the Czech Republic (P305/10/2141) to Z.K. by the Ministry of Education, Youth and Sports of the Czech Republic (LO1220) to P.B. and by IMG institutional support RVO68378050".

VIII CONCLUSIONS

One of the goals of this study was to look into whether BMP signaling is required for axial patterning and for cell fate in the amphioxus embryo. In specific, we wanted to investigate if the BMP positive feedback loop is present in amphioxus similarly as in vertebrates, where it is directly linked to early separation of cell fates leading to the development of a CNS and chorda.

We have studied the role of the *vox* HB gene in brain development. We have specifically addressed the relationship between *vox* and the MHB program. We have used fish *Oryzias latipes* (medaka), an important model in studies of vertebrate development and comparative genomics. We have performed synteny and sequence conservation analysis to validate the identity of the medaka *vox* ortholog. We have described its expression pattern during embryogenesis and placed it into the framework of known brain markers. Using a heat-shock inducible *vox* line of medaka we have shown that *vox* interferes with the GRN during the maintenance stage of developing MHB. Based on our data, we propose a novel function for the *vox* gene as a negative regulator of *fgf8*, *a* pivotal organizer molecule of MHB.

Further, we also sought to identify splicing events in *Pax6* and *Pax2/5/8* classes that influence the paired domain and study the expressions of these alternative spliced transcripts in chordates. We detected three new splicing events in medaka *Pax2* genes, one of which suggests to be highly conserved in Acanthomorphata. We identified a re-occurring splicing variant in medaka and *Danio rerio* (zebrafish) *Pax6* genes using the exon (5a) insertion.

In addition, we have characterized a complement of 18 opsin genes in cubozoan jellyfish *T. cystophora* by whole-genome analysis. Based on phylogenetic, expression and biochemical approaches we assume that rapid lineage- and species-specific duplications of the opsin genes and their following functional diversification promoted evolution of both visual and extraocular photoreception in cubozoans.

IX REFERENCES

Balemans W, Van Hul W (2002) Extracellular regulation of BMP signaling in vertebrates: a cocktail of modulators Developmental biology 250:231-250

Bally-Cuif L, Cholley B, Wassef M (1995) Involvement of Wnt-1 in the formation of the mes/metencephalic boundary Mechanisms of development 53:23-34

Bauer H, Lele Z, Rauch GJ, Geisler R, Hammerschmidt M (2001) The type I serine/threonine kinase receptor Alk8/Lost-a-fin is required for Bmp2b/7 signal transduction during dorsoventral patterning of the zebrafish embryo Development 128:849-858

Belting HG, Wendik B, Lunde K, Leichsenring M, Mossner R, Driever W, Onichtchouk D (2011) Pou5f1 contributes to dorsoventral patterning by positive regulation of vox and modulation of fgf8a expression Developmental biology 356:323-336 doi:10.1016/j.ydbio.2011.05.660

Berry R, Jowitt TA, Garrigue-Antar L, Kadler KE, Baldock C (2010) Structural and functional evidence for a substrate exclusion mechanism in mammalian tolloid like-1 (TLL-1) proteinase FEBS letters 584:657-661 doi:10.1016/j.febslet.2009.12.050

Bier E, De Robertis EM (2015) EMBRYO DEVELOPMENT. BMP gradients: A paradigm for morphogen-mediated developmental patterning Science 348:aaa5838 doi:10.1126/science.aaa5838

Burglin TR (2011) Homeodomain subtypes and functional diversity Sub-cellular biochemistry 52:95-122 doi:10.1007/978-90-481-9069-0_5

Canning CA, Lee L, Irving C, Mason I, Jones CM (2007) Sustained interactive Wnt and FGF signaling is required to maintain isthmic identity Developmental biology 305:276-286 doi:10.1016/j.ydbio.2007.02.009

Cao X, Chen D (2005) The BMP signaling and in vivo bone formation Gene 357:1-8 doi:10.1016/j.gene.2005.06.017

Chen Y, Bhushan A, Vale W (1997) Smad8 mediates the signaling of the ALK-2 [corrected] receptor serine kinase Proceedings of the National Academy of Sciences of the United States of America 94:12938-12943

Chi CL, Martinez S, Wurst W, Martin GR (2003) The isthmic organizer signal FGF8 is required for cell survival in the prospective midbrain and cerebellum Development 130:2633-2644

Chikazu D et al. (2002) Bone morphogenetic protein 2 induces cyclo-oxygenase 2 in osteoblasts via a Cbfal binding site: role in effects of bone morphogenetic protein 2 in vitro and in vivo Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research 17:1430-1440 doi:10.1359/jbmr.2002.17.8.1430

Crossley PH, Martinez S, Martin GR (1996) Midbrain development induced by FGF8 in the chick embryo Nature 380:66-68 doi:10.1038/380066a0

Derynck R, Zhang Y, Feng XH (1998) Smads: transcriptional activators of TGF-beta responses Cell 95:737-740

Dick A et al. (2000) Essential role of Bmp7 (snailhouse) and its prodomain in dorsoventral patterning of the zebrafish embryo Development 127:343-354

Dworkin S, Jane SM (2013) Novel mechanisms that pattern and shape the midbrain-hindbrain boundary Cell Mol Life Sci 70:3365-3374 doi:10.1007/s00018-012-1240-x

Echelard Y, Epstein DJ, St-Jacques B, Shen L, Mohler J, McMahon JA, McMahon AP (1993) Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity Cell 75:1417-1430

Echevarria D, Vieira C, Gimeno L, Martinez S (2003) Neuroepithelial secondary organizers and cell fate specification in the developing brain Brain Res Brain Res Rev 43:179-191

Fekany K et al. (1999) The zebrafish bozozok locus encodes Dharma, a homeodomain protein essential for induction of gastrula organizer and dorsoanterior embryonic structures Development 126:1427-1438

Flores MV, Lam EY, Crosier KE, Crosier PS (2008) Osteogenic transcription factor Runx2 is a maternal determinant of dorsoventral patterning in zebrafish Nat Cell Biol 10:346-352 doi:10.1038/ncb1697

Gao H, Le Y, Wu X, Silberstein LE, Giese RW, Zhu Z (2010) VentX, a novel lymphoid-enhancing factor/T-cell factor-associated transcription repressor, is a putative tumor suppressor Cancer research 70:202-211 doi:10.1158/0008-5472.CAN-09-2668

Gao H, Wu B, Giese R, Zhu Z (2007) Xom interacts with and stimulates transcriptional activity of LEF1/TCFs: implications for ventral cell fate determination during vertebrate embryogenesis Cell Res 17:345-356 doi:10.1038/cr.2007.20

Gao H, Wu X, Sun Y, Zhou S, Silberstein LE, Zhu Z (2012) Suppression of homeobox transcription factor VentX promotes expansion of human hematopoietic stem/multipotent progenitor cells The Journal of biological chemistry 287:29979-29987 doi:10.1074/jbc.M112.383018

Gilardelli CN, Pozzoli O, Sordino P, Matassi G, Cotelli F (2004) Functional and hierarchical interactions among zebrafish vox/vent homeobox genes Dev Dyn 230:494-508 doi:10.1002/dvdy.20073

Glinka A, Wu W, Delius H, Monaghan AP, Blumenstock C, Niehrs C (1998) Dickkopf-1 is a member of a new family of secreted proteins and functions in head induction Nature 391:357-362 doi:10.1038/34848

Glinka A, Wu W, Onichtchouk D, Blumenstock C, Niehrs C (1997) Head induction by simultaneous repression of Bmp and Wnt signalling in Xenopus Nature 389:517-519 doi:10.1038/39092

Gonzalez EM, Fekany-Lee K, Carmany-Rampey A, Erter C, Topczewski J, Wright CV, Solnica-Krezel L (2000) Head and trunk in zebrafish arise via coinhibition of BMP signaling by bozozok and chordino Genes & development 14:3087-3092

Hammerschmidt M et al. (1996) Mutations affecting morphogenesis during gastrulation and tail formation in the zebrafish, Danio rerio Development 123:143-151

Hanai J et al. (1999) Interaction and functional cooperation of PEBP2/CBF with Smads. Synergistic induction of the immunoglobulin germline Calpha promoter The Journal of biological chemistry 274:31577-31582

Hashimoto K, Yokouchi Y, Yamamoto M, Kuroiwa A (1999) Distinct signaling molecules control Hoxa-11 and Hoxa-13 expression in the muscle precursor and mesenchyme of the chick limb bud Development 126:2771-2783

Heldin CH, Miyazono K, ten Dijke P (1997) TGF-beta signalling from cell membrane to nucleus through SMAD proteins Nature 390:465-471 doi:10.1038/37284

Hild M, Dick A, Rauch GJ, Meier A, Bouwmeester T, Haffter P, Hammerschmidt M (1999) The smad5 mutation somitabun blocks Bmp2b signaling during early dorsoventral patterning of the zebrafish embryo Development 126:2149-2159

Hoodless PA, Haerry T, Abdollah S, Stapleton M, O'Connor MB, Attisano L, Wrana JL (1996) MADR1, a MAD-related protein that functions in BMP2 signaling pathways Cell 85:489-500

Houart C, Westerfield M, Wilson SW (1998) A small population of anterior cells patterns the forebrain during zebrafish gastrulation Nature 391:788-792 doi:10.1038/35853

Imai Y, Gates MA, Melby AE, Kimelman D, Schier AF, Talbot WS (2001) The homeobox genes vox and vent are redundant repressors of dorsal fates in zebrafish Development 128:2407-2420

Jessell TM, Sanes JR (2000) Development. The decade of the developing brain Curr Opin Neurobiol 10:599-611

Jesuthasan S, Stahle U (1997) Dynamic microtubules and specification of the zebrafish embryonic axis Current biology : CB 7:31-42

Joyner AL, Liu A, Millet S (2000) Otx2, Gbx2 and Fgf8 interact to position and maintain a mid-hindbrain organizer Curr Opin Cell Biol 12:736-741

Kawabata M, Imamura T, Miyazono K (1998) Signal transduction by bone morphogenetic proteins Cytokine & growth factor reviews 9:49-61

Kawabata M, Imamura T, Miyazono K, Engel ME, Moses HL (1995) Interaction of the transforming growth factor-beta type I receptor with farnesyl-protein transferase-alpha The Journal of biological chemistry 270:29628-29631

Kawahara A, Wilm T, Solnica-Krezel L, Dawid IB (2000a) Antagonistic role of vega1 and bozozok/dharma homeobox genes in organizer formation Proceedings of the National Academy of Sciences of the United States of America 97:12121-12126 doi:10.1073/pnas.97.22.12121

Kawahara A, Wilm T, Solnica-Krezel L, Dawid IB (2000b) Functional interaction of vega2 and goosecoid homeobox genes in zebrafish Genesis 28:58-67

Khokha MK, Yeh J, Grammer TC, Harland RM (2005) Depletion of three BMP antagonists from Spemann's organizer leads to a catastrophic loss of dorsal structures Developmental cell 8:401-411 doi:10.1016/j.devcel.2005.01.013

Kimelman D, Pyati UJ (2005) Bmp signaling: turning a half into a whole Cell 123:982-984 doi:10.1016/j.cell.2005.11.028

Kishimoto Y, Lee KH, Zon L, Hammerschmidt M, Schulte-Merker S (1997) The molecular nature of zebrafish swirl: BMP2 function is essential during early dorsoventral patterning Development 124:4457-4466

Kodjabachian L, Dawid IB, Toyama R (1999) Gastrulation in zebrafish: what mutants teach us Developmental biology 213:231-245 doi:10.1006/dbio.1999.9392

Koenig BB et al. (1994) Characterization and cloning of a receptor for BMP-2 and BMP-4 from NIH 3T3 cells Molecular and cellular biology 14:5961-5974

Koos DS, Ho RK (1999) The nieuwkoid/dharma homeobox gene is essential for bmp2b repression in the zebrafish pregastrula Developmental biology 215:190-207 doi:10.1006/dbio.1999.9479

Kozmikova I, Candiani S, Fabian P, Gurska D, Kozmik Z (2013) Essential role of Bmp signaling and its positive feedback loop in the early cell fate evolution of chordates Developmental biology 382:538-554 doi:10.1016/j.ydbio.2013.07.021

Larabell CA et al. (1997) Establishment of the dorso-ventral axis in Xenopus embryos is presaged by early asymmetries in beta-catenin that are modulated by the Wnt signaling pathway The Journal of cell biology 136:1123-1136

Larroux C, Fahey B, Degnan SM, Adamski M, Rokhsar DS, Degnan BM (2007) The NK homeobox gene cluster predates the origin of Hox genes Current biology : CB 17:706-710 doi:10.1016/j.cub.2007.03.008

Laughon A, Scott MP (1984) Sequence of a Drosophila segmentation gene: protein structure homology with DNA-binding proteins Nature 310:25-31

Leboy P et al. (2001) Smad-Runx interactions during chondrocyte maturation The Journal of bone and joint surgery American volume 83-A Suppl 1:S15-22

Lee KJ, Jessell TM (1999) The specification of dorsal cell fates in the vertebrate central nervous system Annu Rev Neurosci 22:261-294 doi:10.1146/annurev.neuro.22.1.261

Lee KS et al. (2000) Runx2 is a common target of transforming growth factor beta1 and bone morphogenetic protein 2, and cooperation between Runx2 and Smad5 induces osteoblast-specific gene expression in the pluripotent mesenchymal precursor cell line C2C12 Molecular and cellular biology 20:8783-8792

Liu A, Joyner AL (2001) EN and GBX2 play essential roles downstream of FGF8 in patterning the mouse mid/hindbrain region Development 128:181-191

Lun K, Brand M (1998) A series of no isthmus (noi) alleles of the zebrafish pax2.1 gene reveals multiple signaling events in development of the midbrain-hindbrain boundary Development 125:3049-3062

Luyten FP et al. (1989) Purification and partial amino acid sequence of osteogenin, a protein initiating bone differentiation The Journal of biological chemistry 264:13377-13380

Macias-Silva M, Hoodless PA, Tang SJ, Buchwald M, Wrana JL (1998) Specific activation of Smad1 signaling pathways by the BMP7 type I receptor, ALK2 The Journal of biological chemistry 273:25628-25636

Martinez-Barbera JP, Toresson H, Da Rocha S, Krauss S (1997) Cloning and expression of three members of the zebrafish Bmp family: Bmp2a, Bmp2b and Bmp4 Gene 198:53-59

Melby AE, Beach C, Mullins M, Kimelman D (2000) Patterning the early zebrafish by the opposing actions of bozozok and vox/vent Developmental biology 224:275-285 doi:10.1006/dbio.2000.9780

Mizuno T, Yamaha E, Kuroiwa A, Takeda H (1999) Removal of vegetal yolk causes dorsal deficencies and impairs dorsal-inducing ability of the yolk cell in zebrafish Mechanisms of development 81:51-63

Moon JS et al. (2014) Relaxin augments BMP-2-induced osteoblast differentiation and bone formation Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research 29:1586-1596 doi:10.1002/jbmr.2197

Moustakas A, Heldin CH (2002) From mono- to oligo-Smads: the heart of the matter in TGF-beta signal transduction Genes & development 16:1867-1871 doi:10.1101/gad.1016802

Mullins MC (1999) Embryonic axis formation in the zebrafish Methods in cell biology 59:159-178

Mullins MC et al. (1996) Genes establishing dorsoventral pattern formation in the zebrafish embryo: the ventral specifying genes Development 123:81-93

Nguyen VH, Schmid B, Trout J, Connors SA, Ekker M, Mullins MC (1998) Ventral and lateral regions of the zebrafish gastrula, including the neural crest progenitors, are established by a bmp2b/swirl pathway of genes Developmental biology 199:93-110 doi:10.1006/dbio.1998.8927

Nishimura R, Kato Y, Chen D, Harris SE, Mundy GR, Yoneda T (1998) Smad5 and DPC4 are key molecules in mediating BMP-2-induced osteoblastic differentiation of the pluripotent mesenchymal precursor cell line C2C12 The Journal of biological chemistry 273:1872-1879

Nohno T, Ishikawa T, Saito T, Hosokawa K, Noji S, Wolsing DH, Rosenbaum JS (1995) Identification of a human type II receptor for bone morphogenetic protein-4 that forms differential heteromeric complexes with bone morphogenetic protein type I receptors The Journal of biological chemistry 270:22522-22526

Piccolo S, Sasai Y, Lu B, De Robertis EM (1996) Dorsoventral patterning in Xenopus: inhibition of ventral signals by direct binding of chordin to BMP-4 Cell 86:589-598

Raible F, Brand M (2004) Divide et Impera--the midbrain-hindbrain boundary and its organizer Trends Neurosci 27:727-734 doi:10.1016/j.tins.2004.10.003

Ramel MC, Lekven AC (2004) Repression of the vertebrate organizer by Wnt8 is mediated by Vent and Vox Development 131:3991-4000 doi:10.1242/dev.01277

Reddi AH, Reddi A (2009) Bone morphogenetic proteins (BMPs): from morphogens to metabologens Cytokine & growth factor reviews 20:341-342 doi:10.1016/j.cytogfr.2009.10.015

Reifers F, Bohli H, Walsh EC, Crossley PH, Stainier DY, Brand M (1998) Fgf8 is mutated in zebrafish acerebellar (ace) mutants and is required for maintenance of midbrain-hindbrain boundary development and somitogenesis Development 125:2381-2395

Reilly GC, Golden EB, Grasso-Knight G, Leboy PS (2005) Differential effects of ERK and p38 signaling in BMP-2 stimulated hypertrophy of cultured chick sternal chondrocytes Cell communication and signaling : CCS 3:3 doi:10.1186/1478-811X-3-3

Reversade B, Kuroda H, Lee H, Mays A, De Robertis EM (2005) Depletion of Bmp2, Bmp4, Bmp7 and Spemann organizer signals induces massive brain formation in Xenopus embryos Development 132:3381-3392 doi:10.1242/dev.01901

Rhinn M, Brand M (2001) The midbrain--hindbrain boundary organizer Curr Opin Neurobiol 11:34-42

Rhinn M, Lun K, Ahrendt R, Geffarth M, Brand M (2009) Zebrafish gbx1 refines the midbrain-hindbrain boundary border and mediates the Wnt8 posteriorization signal Neural Dev 4:12 doi:10.1186/1749-8104-4-12

Rhinn M, Lun K, Amores A, Yan YL, Postlethwait JH, Brand M (2003) Cloning, expression and relationship of zebrafish gbx1 and gbx2 genes to Fgf signaling Mechanisms of development 120:919-936

Rhinn M, Lun K, Luz M, Werner M, Brand M (2005) Positioning of the midbrain-hindbrain boundary organizer through global posteriorization of the neuroectoderm mediated by Wnt8 signaling Development 132:1261-1272 doi:10.1242/dev.01685

Rowitch DH, McMahon AP (1995) Pax-2 expression in the murine neural plate precedes and encompasses the expression domains of Wnt-1 and En-1 Mechanisms of development 52:3-8

Sasai Y, Lu B, Steinbeisser H, Geissert D, Gont LK, De Robertis EM (1994) Xenopus chordin: a novel dorsalizing factor activated by organizer-specific homeobox genes Cell 79:779-790

Sato T, Joyner AL (2009) The duration of Fgf8 isthmic organizer expression is key to patterning different tectal-isthmo-cerebellum structures Development 136:3617-3626 doi:10.1242/dev.041210

Schier AF (2001) Axis formation and patterning in zebrafish Current opinion in genetics & development 11:393-404

Schmid B, Furthauer M, Connors SA, Trout J, Thisse B, Thisse C, Mullins MC (2000) Equivalent genetic roles for bmp7/snailhouse and bmp2b/swirl in dorsoventral pattern formation Development 127:957-967

Schulte-Merker S, Lee KJ, McMahon AP, Hammerschmidt M (1997) The zebrafish organizer requires chordino Nature 387:862-863 doi:10.1038/43092

Shimizu T, Yamanaka Y, Nojima H, Yabe T, Hibi M, Hirano T (2002) A novel repressor-type homeobox gene, ved, is involved in dharma/bozozok-mediated dorsal organizer formation in zebrafish Mechanisms of development 118:125-138

Sokol SY (2015) Spatial and temporal aspects of Wnt signaling and planar cell polarity during vertebrate embryonic development Seminars in cell & developmental biology 42:78-85 doi:10.1016/j.semcdb.2015.05.002

Solnica-Krezel L, Driever W (2001) The role of the homeodomain protein Bozozok in zebrafish axis formation The International journal of developmental biology 45:299-310

Su CY, Kemp HA, Moens CB (2014) Cerebellar development in the absence of Gbx function in zebrafish Developmental biology 386:181-190 doi:10.1016/j.ydbio.2013.10.026

ten Dijke P et al. (1994) Identification of type I receptors for osteogenic protein-1 and bone morphogenetic protein-4 The Journal of biological chemistry 269:16985-16988

Thisse B et al. (2004) Spatial and temporal expression of the zebrafish genome by large-scale in situ hybridization screening Methods in cell biology 77:505-519

Tuazon FB, Mullins MC (2015) Temporally coordinated signals progressively pattern the anteroposterior and dorsoventral body axes Seminars in cell & developmental biology 42:118-133 doi:10.1016/j.semcdb.2015.06.003

Urist MR (1965) Bone: formation by autoinduction Science 150:893-899

Van Hul E et al. (2002) Localization of the gene causing autosomal dominant osteopetrosis type I to chromosome 11q12-13 Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research 17:1111-1117 doi:10.1359/jbmr.2002.17.6.1111

Varga M, Maegawa S, Bellipanni G, Weinberg ES (2007) Chordin expression, mediated by Nodal and FGF signaling, is restricted by redundant function of two beta-catenins in the zebrafish embryo Mechanisms of development 124:775-791 doi:10.1016/j.mod.2007.05.005

Wang RN et al. (2014) Bone Morphogenetic Protein (BMP) signaling in development and human diseases Genes & Diseases 1:87-105 doi:10.1016/j.gendis.2014.07.005

Wessely O, Kim JI, Geissert D, Tran U, De Robertis EM (2004) Analysis of Spemann organizer formation in Xenopus embryos by cDNA macroarrays Developmental biology 269:552-566 doi:10.1016/j.ydbio.2004.01.018

Wilson SW, Houart C (2004) Early steps in the development of the forebrain Developmental cell 6:167-181

Wittmann DM, Blochl F, Trumbach D, Wurst W, Prakash N, Theis FJ (2009) Spatial analysis of expression patterns predicts genetic interactions at the mid-hindbrain boundary PLoS Comput Biol 5:e1000569 doi:10.1371/journal.pcbi.1000569

Wozney JM (1992) The bone morphogenetic protein family and osteogenesis Molecular reproduction and development 32:160-167 doi:10.1002/mrd.1080320212

Wozney JM et al. (1988) Novel regulators of bone formation: molecular clones and activities Science 242:1528-1534

Wu X, Gao H, Bleday R, Zhu Z (2014) Homeobox transcription factor VentX regulates differentiation and maturation of human dendritic cells The Journal of biological chemistry 289:14633-14643 doi:10.1074/jbc.M113.509158

Wu X, Gao H, Ke W, Giese RW, Zhu Z (2011a) The homeobox transcription factor VentX controls human macrophage terminal differentiation and proinflammatory activation The Journal of clinical investigation 121:2599-2613 doi:10.1172/JCI45556

Wu X, Gao H, Ke W, Hager M, Xiao S, Freeman MR, Zhu Z (2011b) VentX trans-activates p53 and p16ink4a to regulate cellular senescence The Journal of biological chemistry 286:12693-12701 doi:10.1074/jbc.M110.206078

Wurst W, Bally-Cuif L (2001) Neural plate patterning: upstream and downstream of the isthmic organizer Nat Rev Neurosci 2:99-108 doi:10.1038/35053516

Yamanaka Y et al. (1998) A novel homeobox gene, dharma, can induce the organizer in a non-cell-autonomous manner Genes & development 12:2345-2353

Yamashita H et al. (1995) Osteogenic protein-1 binds to activin type II receptors and induces certain activin-like effects The Journal of cell biology 130:217-226

Zhao J, Lambert G, Meijer AH, Rosa FM (2013) The transcription factor Vox represses endoderm development by interacting with Casanova and Pou2 Development 140:1090-1099 doi:10.1242/dev.082008

Zhao M, Qiao M, Oyajobi BO, Mundy GR, Chen D (2003) E3 ubiquitin ligase Smurf1 mediates core-binding factor alpha1/Runx2 degradation and plays a specific role in osteoblast differentiation The Journal of biological chemistry 278:27939-27944 doi:10.1074/jbc.M304132200

Zhong YF, Holland PW (2011) The dynamics of vertebrate homeobox gene evolution: gain and loss of genes in mouse and human lineages BMC evolutionary biology 11:169 doi:10.1186/1471-2148-11-169

Zimmerman LB, De Jesus-Escobar JM, Harland RM (1996) The Spemann organizer signal noggin binds and inactivates bone morphogenetic protein 4 Cell 86:599-606